

Similar to the constructions of the vectors $\begin{pmatrix} (1) \\ T \end{pmatrix}$, $\begin{pmatrix} (1) \\ 1,T \end{pmatrix}$, $\begin{pmatrix} (1) \\ 2,T \end{pmatrix}$, and $\begin{pmatrix} (1) \\ T \end{pmatrix}$, we define $\begin{pmatrix} (1) \\ i,T \end{pmatrix}$, $\hat{\begin{pmatrix} (1) \\ 1,T \end{pmatrix}}$, $\hat{\begin{pmatrix} (1) \\ 2,T \end{pmatrix}}$, and $\hat{\begin{pmatrix} (1) \\ T \end{pmatrix}}$ as restricting the vectors $\begin{pmatrix} (1) \\ i \end{pmatrix}$, $\hat{\begin{pmatrix} (1) \\ 1 \end{pmatrix}}$, $\hat{\begin{pmatrix} (1) \\ 2 \end{pmatrix}}$, and $\hat{\begin{pmatrix} (1) \\ T \end{pmatrix}}$ to the discriminant set T .

Given index sets T and N , we define several sample covariance matrices as

$$\begin{aligned} S^{(1)} &= f(n_1 - 1)S_1^{(1)} + (n_2 - 1)S_2^{(1)}g=(n - 2); \\ S_{TT}^{(1)} &= f(n_1 - 1)S_{1,TT}^{(1)} + (n_2 - 1)S_{2,TT}^{(1)}g=(n - 2); \\ S_{NT}^{(1)} &= f(n_1 - 1)S_{1,NT}^{(1)} + (n_2 - 1)S_{2,NT}^{(1)}g=(n - 2); \end{aligned}$$

where

$$\begin{aligned} S_1^{(1)} &= \prod_{i \in H_1} \begin{pmatrix} (1) \\ i \end{pmatrix} \begin{pmatrix} (1) \\ i \end{pmatrix}' = (n_1 - 1); \\ S_2^{(1)} &= \prod_{i \in H_2} \begin{pmatrix} (1) \\ i \end{pmatrix} \begin{pmatrix} (1) \\ i \end{pmatrix}' = (n_2 - 1); \\ S_{1,TT}^{(1)} &= \prod_{i \in H_1} \begin{pmatrix} (1) \\ i,T \end{pmatrix} \begin{pmatrix} (1) \\ i,T \end{pmatrix}' = (n_1 - 1); \\ S_{2,TT}^{(1)} &= \prod_{i \in H_2} \begin{pmatrix} (1) \\ i,T \end{pmatrix} \begin{pmatrix} (1) \\ i,T \end{pmatrix}' = (n_2 - 1); \\ S_{1,NT}^{(1)} &= \prod_{i \in H_1} \begin{pmatrix} (1) \\ i,N \end{pmatrix} \begin{pmatrix} (1) \\ i,T \end{pmatrix}' = (n_1 - 1); \\ S_{2,NT}^{(1)} &= \prod_{i \in H_2} \begin{pmatrix} (1) \\ i,N \end{pmatrix} \begin{pmatrix} (1) \\ i,T \end{pmatrix}' = (n_2 - 1); \end{aligned}$$

Similar to the definitions of $\begin{pmatrix} (1) \\ 1 \end{pmatrix}$, $\begin{pmatrix} (1) \\ 2 \end{pmatrix}$, $\begin{pmatrix} (1) \\ 1,T \end{pmatrix}$, $\begin{pmatrix} (1) \\ 2,T \end{pmatrix}$, and $\begin{pmatrix} (1) \\ T \end{pmatrix}$, we denote for any $\ell = 1, 2$,

$$\begin{aligned} \tilde{\ell}^{(2)} &= E(\tilde{\ell}^{(2)} j Y = \cdot) = (\tilde{\ell}_1^{(2)'}; \dots; \tilde{\ell}_{p_n}^{(2)'})'; \\ \tilde{\ell}_j^{(2)} &= E(\tilde{\ell}_j^{(2)} j Y = \cdot) = (\ell_{j,sn+1}; \ell_{j,sn+2}; \dots)' \in \mathbb{R}^\infty; \quad j = 1; \dots; p_n; \\ \tilde{\ell}_{\ell,T}^{(2)} &: \text{formed by stacking } \tilde{\ell}_j^{(2)} : j \in T \text{ in a column,} \\ \tilde{\ell}^{(2)} &= \begin{pmatrix} (2) \\ 2 \end{pmatrix} \begin{pmatrix} (2) \\ 1 \end{pmatrix}; \quad \tilde{\ell}_T^{(2)} = \begin{pmatrix} (2) \\ 2,T \end{pmatrix} \begin{pmatrix} (2) \\ 1,T \end{pmatrix}. \end{aligned}$$

Similar to the constructions of $\mathcal{R}^{(1)}$ and $\mathcal{R}_T^{(1)}$, we denote $\mathcal{R}^{(1)}$, $\mathcal{R}_T^{(1)}$, $\mathcal{R}^{(2)}$, and $\mathcal{R}_T^{(2)}$ as

$$\begin{aligned} \mathcal{R}^{(1)} &= (\mathcal{R}_1^{(1)'} \cdots \mathcal{R}_{p_n}^{(1)'})' \text{ with each } \mathcal{R}_j^{(1)} = (\mathcal{R}_{j_1} \cdots \mathcal{R}_{j_{s_n}})' ; \\ \mathcal{R}_T^{(1)} &: \text{ formed by stacking } \mathcal{R}_j^{(1)} : j \geq Tg \text{ in a column,} \\ \mathcal{R}^{(2)} &= (\mathcal{R}_1^{(2)'} \cdots \mathcal{R}_{p_n}^{(2)'})' \text{ with each } \mathcal{R}_j^{(2)} = (\mathcal{R}_{j, s_n+1} \cdots \mathcal{R}_{j, s_n+2} \cdots)' ; \\ \mathcal{R}_T^{(2)} &: \text{ formed by stacking } \mathcal{R}_j^{(2)} : j \geq Tg \text{ in a column.} \end{aligned}$$

In the next section, we present the proofs of the main results, Theorems 1-2 and Corollary 1.

2 Proofs of Theorems 1-2 and Corollary 1

Proof of Theorem 1: Under conditions (A1) and (A2), property (i) holds directly from Lemma 1. To show property (ii), first note that

$$\begin{aligned} \Delta &= (\mathcal{R}'_{T^*} \Sigma_{T^* T^*} \mathcal{R}_{T^*})^{1/2} = f(\Lambda_{T^*}^{1/2} \mathcal{R}'_{T^*})' (\Lambda_{T^*}^{\dagger 1/2} \Sigma_{T^* T^*} \Lambda_{T^*}^{\dagger 1/2}) (\Lambda_{T^*}^{1/2} \mathcal{R}_{T^*}) g^{1/2} \\ &= c_1^{1/2} k \Lambda_{T^*}^{1/2} \mathcal{R}'_{T^*} k_2 = c_1^{1/2} \left(\prod_{j \in T^*} \prod_{k=1}^{\infty} \lambda_{jk}^{*2} \right)^{1/2}. \end{aligned}$$

Together with condition (A3), it can be seen that

$$\Delta \asymp 1; \text{ as } n \rightarrow \infty. \quad (1)$$

Hence, property (ii) holds from (6) in the main paper and (1). To show property (iii), first note that

$$\Delta^{(1)} = f_1 + o(r_n^{-1}) + o(r_n^{-1/2} \lambda_n^{1/2}) g \Delta \asymp 1; \quad (2)$$

by Lemma 1 and (1). Moreover, by definition, it is not hard to verify that

$$R(\mathcal{R}^{(1)}) = R^{(1)}(\mathcal{R}^{(1)}) = (I + \lambda_2 \Omega_1) (I + \lambda_2 \Omega_2)^{-1} \Omega_3; \quad (3)$$

where

$$\begin{aligned}\Omega_1 &= \Phi(\Delta=2 + \log(\lambda_1 = \lambda_2) = \Delta) = \Phi(\Delta=2 + \log(\lambda_2 = \lambda_1) = \Delta); \\ \Omega_2 &= \Phi(\Delta^{(1)}=2 + \log(\lambda_1 = \lambda_2) = \Delta^{(1)}) = \Phi(\Delta^{(1)}=2 + \log(\lambda_2 = \lambda_1) = \Delta^{(1)}); \\ \Omega_3 &= \Phi(\Delta=2 + \log(\lambda_2 = \lambda_1) = \Delta) = \Phi(\Delta^{(1)}=2 + \log(\lambda_2 = \lambda_1) = \Delta^{(1)});\end{aligned}$$

For the term Ω_1 , it can be rewritten as

$$\Omega_1 = \Phi(\%_n^{1/2}(1 + \#_n)) = \Phi(\%_n^{1/2}); \quad (4)$$

where $\%_n = f\Delta=2 \log(\lambda_2 = \lambda_1) = \Delta g^2$ and $\#_n = 4\log(\lambda_2 = \lambda_1) = f\Delta^2 - 2\log(\lambda_2 = \lambda_1)g$. Since $\%_n \rightarrow 1$ and $\%_n \#_n \rightarrow \log(\lambda_2 = \lambda_1)$ under (1), we immediately conclude that

$$\Omega_1 \rightarrow \lambda_2 = \lambda_1; \quad (5)$$

by applying Lemma 1 of [Shao et al. \(2011\)](#) to (4). Similar argument leads to

$$\Omega_2 \rightarrow \lambda_2 = \lambda_1; \quad (6)$$

For the term Ω_3 , it can be expressed as

$$\Omega_3 = \Phi(\%_n^{1/2}(1 + \tilde{\#}_n)) = \Phi(\%_n^{1/2}); \quad (7)$$

where $\%_n = f\Delta^{(1)}=2 \log(\lambda_2 = \lambda_1) = \Delta^{(1)}g^2$ and $\tilde{\#}_n = [f\Delta\Delta^{(1)}+2\log(\lambda_2 = \lambda_1)g(\Delta - \Delta^{(1)})] = f\Delta\Delta^{(1)2} - 2\log(\lambda_2 = \lambda_1)\Delta g$. Based on (2) and (A3), one can show that

$$\%_n \rightarrow 1; \quad \%_n \tilde{\#}_n \rightarrow 0;$$

Together with (7) and Lemma 1 of [Shao et al. \(2011\)](#), it can be concluded that

$$\Omega_3 \rightarrow 1;$$

Together with (3), (5) and (6), we have $R(\lambda^*) = R^o(\lambda^{(1)}) \rightarrow 1$, which completes the proof. \square

Remark: Although not part of the proof, it is important to justify that the ideal classifier in (3) of the main article is really the optimal rule. By definition, we have

$$jY = 1 \quad N(\mu_1; \Sigma); \quad jY = 2 \quad N(\mu_2; \Sigma);$$

which implies

$$\Sigma^{\dagger 1/2} jY = 1 \quad N(\Sigma^{\dagger 1/2} \mu_1; I); \quad \Sigma^{\dagger 1/2} jY = 2 \quad N(\Sigma^{\dagger 1/2} \mu_2; I);$$

Therefore, the conditional density functions of $Z = \Sigma^{\dagger 1/2} z$ take the form:

$$f_z(z|Y = i) \propto \exp\left\{-\frac{1}{2}(z - \Sigma^{\dagger 1/2} \mu_i)'(z - \Sigma^{\dagger 1/2} \mu_i)\right\} g_i; \quad \text{for } i = 1, 2;$$

By change of variables, the conditional density functions of $z = \Sigma^{-1/2} Z$

where \tilde{V}_T is defined in (16) of the main paper and

$$\begin{aligned} \tilde{V}_T &= f n_1 n_2 n^{-1} (n-2)^{-1} g^{\wedge(1)} + \frac{f n_1 n_2 n^{-1} (n-2)^{-1} g^{\wedge(1)'}}{n \hat{\Lambda}_T^{(1)'}} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\frac{(1)}{T}) g + 1 + \\ & f n_1 n_2 n^{-1} (n-2)^{-1} g^{\wedge(1)'}}{n \hat{\Lambda}_T^{(1)'}} S_{TT}^{(1)-1} \frac{\wedge(1)}{T}^{-1} S_{TT}^{(1)-1} \frac{\wedge(1)}{T} \frac{f n_1 n_2 n^{-1} (n-2)^{-1} g^{\wedge(1)'}}{n \hat{\Lambda}_T^{(1)'}} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\frac{(1)}{T}): \end{aligned}$$

To prove property (i), based on (8), (9) and the Karush-Kuhn-Tucker conditions, it is sufficient to show that there exist positive constants $c_5; c_6 > 0$ such that

$$\begin{aligned} P \quad & f n_1 n_2 n^{-1} (n-2)^{-1} g^{\wedge(1)} \quad f S_{TT}^{(1)} + n_1 n_2 n^{-1} (n-2)^{-1} \frac{\wedge(1)}{T} \frac{\wedge(1)'}{T} g \tilde{V}_T = \\ & \frac{f n_1 n_2 n^{-1} (n-2)^{-1} g^{\wedge(1)'}}{n \hat{\Lambda}_T^{(1)'}} \text{sgn}(\tilde{V}_T) \quad 1 \quad c_5 [f(\rho_n \quad q_n) S_n g^{-1} + (q_n S_n)^{-1} + f \log(n) g^{-1} + \\ & \exp(-n-1) + \exp(-n-2)]; \end{aligned} \quad (10)$$

and

$$\begin{aligned} P \quad & \hat{\Lambda}_N^{(1)-1/2} f n_1 n_2 n^{-1} (n-2)^{-1} g^{\wedge(1)} \quad f S_{NT}^{(1)} + n_1 n_2 n^{-1} (n-2)^{-1} \frac{\wedge(1)}{N} \frac{\wedge(1)'}{T} g \\ & \tilde{V}_T \quad \infty \quad n \quad 1 \quad c_6 [f(\rho_n \quad q_n) S_n g^{-1} + (q_n S_n)^{-1} + f \log(n) g^{-1} + \\ & \exp(-n-1) + \exp(-n-2)]; \end{aligned} \quad (11)$$

Note that the random quantity $S_{NT}^{(1)}$ can be expressed as $S_{NT}^{(1)} = f(n_1-1)S_{1,NT}^{(1)} + (n_2-1)S_{2,NT}^{(1)}g=(n-2)$, where $S_{1,NT}^{(1)} = \mathbb{P}_{i \in H_1} (\frac{(1)}{i,N} \quad \frac{\wedge(1)}{1,N}) (\frac{(1)}{i,T} \quad \frac{\wedge(1)}{1,T})' = (n_1-1)$ and $S_{2,NT}^{(1)} = \mathbb{P}_{i \in H_2} (\frac{(1)}{i,N} \quad \frac{\wedge(1)}{2,N}) (\frac{(1)}{i,T} \quad \frac{\wedge(1)}{2,T})' = (n_2-1)$. Since \tilde{V}_T is the solution to the convex optimization problem specified in Lemma 2, the first order condition together with Lemma 11 yields (10) immediately. To show (11), we first note that

$$\begin{aligned} & \hat{\Lambda}_N^{(1)-1/2} f n_1 n_2 n^{-1} (n-2)^{-1} g^{\wedge(1)} \quad f S_{NT}^{(1)} + n_1 n_2 n^{-1} (n-2)^{-1} \\ & \frac{\wedge(1)}{N} \frac{\wedge(1)'}{T} g \tilde{V}_T \quad \infty \quad 1 + k \hat{\Lambda}_N^{(1)-1/2} \Lambda_N^{(1)1/2} \quad l_{(p_n-q_n)S_n} k_{\max} \quad k \Psi k_{\infty}; \end{aligned} \quad (12)$$

where $\Psi = \Lambda_N^{(1)-1/2} f S_{NT}^{(1)} + n_1 n_2 n^{-1} (n-2)^{-1} \frac{\wedge(1)}{N} \frac{\wedge(1)'}{T} g \tilde{V}_T \quad f n_1 n_2 n^{-1} (n-2)^{-1} g^{\wedge(1)}$. By

definition, conditional on any nonempty set $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$\begin{aligned} & (n-2)\Lambda^{(1)-1/2} \mathcal{S}^{(1)} \Lambda^{(1)-1/2} fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \\ & \text{Wishart}(n-2, j\Lambda^{(1)-1/2} \Sigma^{(1)} \Lambda^{(1)-1/2}); \end{aligned} \quad (13)$$

where the set $\mathcal{M}_n = \{f_{1=2} \dots f_{n-1=n-3} \setminus f_{2=2} \dots f_{n-2=2}\}$ is defined in Lemma 3. Moreover, conditional on any nonempty set $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$(n-2)\Lambda^{(1)-1/2} \mathcal{S}^{(1)} \Lambda^{(1)-1/2} \text{? } \wedge^{(1)};$$

where the symbol ? means independent of. Together with (13), it can be concluded that there exists a collection $fZ_l g_{l=1}^{n-2}$ of $n-2$ random vectors in $\mathbb{R}^{pn \times sn}$ satisfying (14) to (16) as follows.

$$(n-2)\Lambda^{(1)-1/2} \mathcal{S}^{(1)} \Lambda^{(1)-1/2} = \prod_{l=1}^{n-2} Z_l Z_l' \quad (14)$$

Conditional on any nonempty set $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$fZ_l g_{l=1}^{n-2} \text{? } \wedge^{(1)}; \quad (15)$$

Conditional on any nonempty set $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$Z_l fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \stackrel{i.i.d}{\sim} \mathcal{N}(0; \Lambda^{(1)-1/2} \Sigma^{(1)} \Lambda^{(1)-1/2}); \quad l = 1; \dots; n-2; \quad (16)$$

For each $l = 1; \dots; n-2$, we write the vector $Z_l = (\tilde{Z}'_{l1}; \dots; \tilde{Z}'_{lpn})' \in \mathbb{R}^{pn \times sn}$ with sub-vectors $\tilde{Z}'_{lj} = (Z_{lj1}; \dots; Z_{ljsn})' \in \mathbb{R}^{sn}$. Similarly, for each $l = 1; \dots; n-2$, we let $Z_{l,T} = (\tilde{Z}'_{l1}; \dots; \tilde{Z}'_{lqn})' \in \mathbb{R}^{qn \times sn}$ and $Z_{l,N} = (\tilde{Z}'_{l,qn+1}; \dots; \tilde{Z}'_{l,pn})' \in \mathbb{R}^{(pn-qn) \times sn}$. Accordingly, we denote

$$\begin{aligned} Z_T &= [Z_{1,T}; \dots; Z_{n-2,T}] \in \mathbb{R}^{qn \times (n-2)}; \\ Z_N &= [Z_{1,N}; \dots; Z_{n-2,N}] \in \mathbb{R}^{(pn-qn) \times (n-2)}; \\ Z &= [Z_T; Z_N]' = [Z_1; \dots; Z_{n-2}] \in \mathbb{R}^{pn \times (n-2)}. \end{aligned} \quad (17)$$

Note that (18) entails that conditional on any nonempty set $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$\begin{aligned} \hat{\wedge}_T^{(1)} &? fZ_T; Z_N \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T g; \\ \hat{\wedge}_T^{(1)} &? Z_N \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T; \\ \hat{\wedge}_T^{(1)} &? Z_T; \end{aligned} \quad (23)$$

Piecing (22) and (23) together yields that conditional on any nonempty set $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$f\hat{\wedge}_T^{(1)}; Z_T g ? Z_N \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T. \quad (24)$$

In a similar fashion, the quantity $\Lambda_N^{(1)-1/2} \hat{\wedge}_N^{(1)}$ can be decomposed into

$$\Lambda_N^{(1)-1/2} \hat{\wedge}_N^{(1)} = \Lambda_N^{(1)-1/2} \left(\hat{\wedge}_N^{(1)} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \hat{\wedge}_T^{(1)} \right) + \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \hat{\wedge}_T^{(1)}. \quad (25)$$

It is not difficult to verify that conditional on any nonempty set $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$\begin{array}{ccc} \begin{array}{c} 2 \\ 6 \\ \circ \\ \circ \\ \circ \\ \circ \\ \circ \\ 4 \end{array} \Lambda_N^{(1)-1/2} \left(\hat{\wedge}_N^{(1)} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \hat{\wedge}_T^{(1)} \right) & \begin{array}{c} 3 \\ 7 \\ 7 \\ 7 \\ 7 \\ 5 \end{array} fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n & \begin{array}{c} 2 \\ 6 \\ \circ \\ \circ \\ \circ \\ \circ \\ \circ \\ 4 \end{array} \Lambda_N^{(1)-1/2} \left(\hat{\wedge}_N^{(1)} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \hat{\wedge}_T^{(1)} \right) \\ & & \begin{array}{c} 2 \\ 7 \\ 7 \\ 7 \\ 7 \\ 5 \end{array} \Lambda_T^{(1)-1/2} \hat{\wedge}_T^{(1)} \end{array} \quad (26)$$

$$\begin{array}{ccc} \begin{array}{c} 2 \\ 6 \\ \circ \\ \circ \\ \circ \\ \circ \\ 4 \end{array} \Lambda_N^{(1)-1/2} \left(\Sigma_{NN}^{(1)} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TN}^{(1)} \right) \Lambda_N^{(1)-1/2} & \begin{array}{c} 3 \\ 7 \\ 7 \\ 7 \\ 7 \\ 5 \end{array} 0_{(p_n - q_n) s_n \times q_n s_n} & \begin{array}{c} 2 \\ 6 \\ \circ \\ \circ \\ \circ \\ \circ \\ 4 \end{array} \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} \\ n n_1^{-1} n_2^{-1} & & \end{array}$$

which further entails that conditional on any nonempty set $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$\hat{\wedge}_T^{(1)} ? \hat{\wedge}_N^{(1)} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \hat{\wedge}_T^{(1)}. \quad (27)$$

Based on (18), it is seen that conditional on any nonempty set $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$\begin{aligned} Z_T & \ ? \ f_T^{(1)}; \hat{\Lambda}_N^{(1)} \ \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)} g; \\ Z_T & \ ? \ \hat{\Lambda}_N^{(1)} \ \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)}; \\ Z_T & \ ? \ \hat{\Lambda}_T^{(1)}; \end{aligned} \tag{28}$$

Together with (27) yields that conditional on any nonempty set $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$f_T^{(1)}; Z_T g \ ? \ \hat{\Lambda}_N^{(1)} \ \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)}. \tag{29}$$

Moreover, using (19) and (25), elementary algebra yields that

$$\Psi = \Pi_1 \ \Pi_2 \ \Pi_3 \ \Pi_4; \tag{30}$$

with

$$\begin{aligned} \Pi_1 &= (n \ 2)^{-1} (Z_N \ \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T) Z_T' \Lambda_T^{(1)1/2} \hat{V}_T; \\ \Pi_2 &= \hat{\#} \Lambda_N^{(1)-1/2} (\hat{\Lambda}_N^{(1)} \ \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)}); \\ \Pi_3 &= \ _n \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} (\Lambda_T^{(1)-1/2} \hat{\Lambda}_T^{(1)1/2} \ l_{q_n s_n}) \text{sgn}(\ \hat{\Lambda}_T^{(1)}); \\ \Pi_4 &= \ _n \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\ \hat{\Lambda}_T^{(1)}); \end{aligned}$$

where

$$\begin{aligned} \hat{\#} &= f n_1 n_2 n^{-1} (n \ 2)^{-1} g f 1 + \ _n \hat{\Lambda}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\ \hat{\Lambda}_T^{(1)}) g \\ & \ 1 + f n_1 n_2 n^{-1} (n \ 2)^{-1} g \hat{\Lambda}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)-1}; \end{aligned}$$

Similar arguments as in the proof of Lemma 11 indicates that there exist universal constants $c_7 > 0$ and $c_9 > c_8 > 0$ such that with probability at least $1 - c_7 [(q_n s_n)^{-1} + f \log(n) g^{-1} + \exp(-n_1=12) + \exp(-n_2=12)]$,

$$c_8 f \ \hat{\Lambda}_T^{(1)'} \Sigma_{TT}^{(1)-1} \ \hat{\Lambda}_T^{(1)} g^{-1} \ \hat{\#} \ c_9 f \ \hat{\Lambda}_T^{(1)'} \Sigma_{TT}^{(1)-1} \ \hat{\Lambda}_T^{(1)} g^{-1}. \tag{31}$$

For the term Π_1 , it can be decomposed into

$$\Pi_1 = \Upsilon_1 + \Upsilon_2; \quad (32)$$

where

$$\begin{aligned} \Upsilon_1 &= \hat{\#}(n-2)^{-1} (Z_N \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T) Z_T' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\wedge}_T^{(1)}; \\ \Upsilon_2 &= {}_n(n-2)^{-1} (Z_N \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T) Z_T' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\hat{\wedge}_T^{(1)}); \end{aligned}$$

At this point, we denote $f e_j g_{j=1}^{(p_n - q_n) s_n}$ as the standard basis in $\mathbb{R}^{(p_n - q_n) s_n}$. Moreover, according to (20), (21) and (24), it can be deduced that conditional on any nonempty set

$$f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_T^{(1)}; Z_T g \text{ and for any } j \in (\rho_n - q_n) S_n,$$

$$\begin{aligned} (Z_N \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} Z_T)' e_j f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_T^{(1)}; Z_T g \\ N(0_{(n-2) \times 1}; f e_j' \Lambda_N^{(1)-1/2} (\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TN}^{(1)}) \Lambda_N^{(1)-1/2} e_j g_{l_{n-2}}); \end{aligned}$$

which implies that conditional on any nonempty set $f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_T^{(1)}; Z_T g$ and for any $j \in (\rho_n - q_n) S_n$,

$$e_j' \Upsilon_1 f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_T^{(1)}; Z_T g \quad N(0; \Gamma_j);$$

with each

$$\begin{aligned} \Gamma_j &= \hat{\#}^2 (n-2)^{-1} f e_j' \Lambda_N^{(1)-1/2} (\Sigma_{NN}^{(1)} - \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TN}^{(1)}) \Lambda_N^{(1)-1/2} e_j g_T^{(1)'} S_{TT}^{(1)-1} \hat{\wedge}_T^{(1)} \\ &\quad \hat{\#}^2 (n-2)^{-1} \hat{\wedge}_T^{(1)'} S_{TT}^{(1)-1} \hat{\wedge}_T^{(1)}. \end{aligned}$$

Together with the maximal inequality, we have that for any $t > 0$,

$$\begin{aligned} P(K \Upsilon_1 k_\infty > t | f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_T^{(1)}; Z_T g) \\ \leq 2(\rho_n - q_n) S_n \exp \left\{ -4^{-1} \hat{\#}^{-2} \hat{f}_T^{(1)'} S_{TT}^{(1)-1} \hat{\wedge}_T^{(1)} g^{-1} n t^2 \right\}; \end{aligned}$$

with each

$$\begin{aligned} \Xi_j &= \frac{2}{n} (n-2)^{-1} \bar{f} e_j' \Lambda_N^{(1)-1/2} (\Sigma_{NN}^{(1)} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TN}^{(1)}) \Lambda_N^{(1)-1/2} e_j g \bar{f} \text{sgn}(\frac{(1)}{T})' \\ &\quad \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\frac{(1)}{T}) g \\ &\quad \frac{2}{n} (n-2)^{-1} \bar{f} \text{sgn}(\frac{(1)}{T})' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\frac{(1)}{T}) g: \end{aligned}$$

Together with maximal inequality, we have that for any $t \geq 0$,

$$\begin{aligned} P \leq & k\Upsilon_2 k_\infty \quad t \bar{f} Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_T^{(1)}; Z_T g \\ & 2(\rho_n - q_n) S_n \exp \left(4^{-1} \frac{2}{n} \bar{f} \text{sgn}(\frac{(1)}{T})' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\frac{(1)}{T}) g^{-1} n t^2 \right): \end{aligned}$$

Setting $t = [8 \frac{2}{n} \bar{f} \text{sgn}(\frac{(1)}{T})' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\frac{(1)}{T}) g \log \bar{f}(\rho_n - q_n) S_n g = n]^{1/2}$ in the above inequality yields

$$\begin{aligned} P &\leq k\Upsilon_2 k_\infty \quad [8 \frac{2}{n} \bar{f} \text{sgn}(\frac{(1)}{T})' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\frac{(1)}{T}) g]^{1/2} \\ &\quad [\log \bar{f}(\rho_n - q_n) S_n g = n]^{1/2} \bar{f} Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_T^{(1)}; Z_T g \\ &\quad 1 - 2 \bar{f}(\rho_n - q_n) S_n g^{-1}: \end{aligned}$$

Together with similar reasoning as in (34), one has

$$\begin{aligned} P &\leq k\Upsilon_2 k_\infty \quad [8 \frac{2}{n} \bar{f} \text{sgn}(\frac{(1)}{T})' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\frac{(1)}{T}) g]^{1/2} \\ &\quad [\log \bar{f}(\rho_n - q_n) S_n g = n]^{1/2} \\ &\quad 1 - c_{13} [\bar{f}(\rho_n - q_n) S_n g^{-1} + \exp(-n_1=12) + \exp(-n_2=12)]; \end{aligned}$$

for some universal constant $c_{13} > 0$. Then, it follows from the above inequality and Lemma 9 that there exist universal constants $c_{14}, c_{15} > 0$ such that with probability at least $1 - c_{14} [\bar{f}(\rho_n - q_n) S_n g^{-1} + (q_n S_n)^{-1} + \bar{f} \log(n) g^{-1} + \exp(-n_1=12) + \exp(-n_2=12)]$,

$$k\Upsilon_2 k_\infty \leq c_{15} [q_n S_n \log \bar{f}(\rho_n - q_n) S_n g = n]^{1/2} n:$$

For the term Π_3 , it follows from condition (C2) and Lemma 5 that there exist universal constants $c_{21}, c_{22} > 0$ such that with probability at least $1 - c_{21} \bar{f}(q_n s_n)^{-1} + \exp(-n_1/12) + \exp(-n_2/12)g$, we have $k\Pi_3 k_\infty \leq c_{22} \bar{f} q_n s_n \log(q_n s_n) = n g^{1/2}$. Together with (37), (36) and (30), there exists a universal constant $c_{23} > 0$ such that with probability at least $1 - c_{23} [\bar{f}(\rho_n - q_n) s_n g^{-1} + (q_n s_n)^{-1} + \bar{f} \log(n) g^{-1} + \exp(-n_1/12) + \exp(-n_2/12)]$, we have $k\Psi k_\infty \leq (1 - c_{23}) n$. Together with (12) and Lemma 6, the assertion (11) holds trivially, which completes the proof of property (i). To show property (iii), we recall that $\tilde{V} = (\tilde{V}_T; 0) \in \mathbb{R}^{p n s n}$, where \tilde{V}_T is defined in Lemma 2. Together with (9), we have that there exists a universal constants $c_{24} > 0$ such that with probability at least $1 - c_{24} \bar{f}(q_n s_n)^{-1} + \bar{f} \log(n) g^{-1} + \exp(-n_1/12) + \exp(-n_2/12)g$,

$$R^\circ(\tilde{V}) = R^\circ(\tilde{V}) = {}_1\Omega_1 + {}_2\Omega_2; \quad (38)$$

where

$$\begin{aligned} \Omega_1 &= \Phi \quad \tilde{V}_T \begin{pmatrix} \hat{\cdot}^{(1)} \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}_{1,T} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}_{1,T} \quad 2^{-1} \tilde{V}_T \hat{\cdot}^{(1)}_T + \bar{f} \tilde{V}_T \mathcal{S}_{TT}^{(1)} \tilde{V}_T g \bar{f} \hat{\cdot}^{(1)}_T g^{-1} \bar{f} \log(n_2 = n_1) g \\ &\quad \bar{f} \tilde{V}_T \Sigma_{TT}^{(1)} \tilde{V}_T g^{-1/2} \quad ; \\ \Omega_2 &= \Phi \quad \tilde{V}_T \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}_{2,T} \hat{\cdot}^{(1)}_{2,T} \quad 2^{-1} \tilde{V}_T \hat{\cdot}^{(1)}_T - \bar{f} \tilde{V}_T \mathcal{S}_{TT}^{(1)} \tilde{V}_T g \bar{f} \hat{\cdot}^{(1)}_T g^{-1} \bar{f} \log(n_2 = n_1) g \\ &\quad \bar{f} \tilde{V}_T \Sigma_{TT}^{(1)} \tilde{V}_T g^{-1/2} \quad ; \end{aligned}$$

Also recalling from (11) of the main paper that

$$R^\circ(\cdot^{(1)}) = {}_1\Omega_1^* + {}_2\Omega_2^*; \quad (39)$$

with

$$\begin{aligned} \Omega_1^* &= \Phi \quad 2^{-1} \bar{f} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}'_T \Sigma_{TT}^{(1)-1} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}_T g^{1/2} + \log(n_2 = n_1) \bar{f} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}'_T \Sigma_{TT}^{(1)-1} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}_T g^{-1/2} \quad ; \\ \Omega_2^* &= \Phi \quad 2^{-1} \bar{f} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}'_T \Sigma_{TT}^{(1)-1} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}_T g^{1/2} - \log(n_2 = n_1) \bar{f} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}'_T \Sigma_{TT}^{(1)-1} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}_T g^{-1/2} \quad ; \end{aligned}$$

We denote a_n , b_n , X_n and U_n as

$$a_n = 4^{-1} \mathbf{1}' \Sigma_{TT}^{(1)-1} \mathbf{1}; \quad b_n = \log(2) \mathbf{1}' \Sigma_{TT}^{(1)-1} \mathbf{1} g^{-1/2};$$

$$X_n = \mathbf{1}' \tilde{V}_T \hat{\Sigma}_{1,T}^{(1)} + \tilde{V}_T \hat{\Sigma}_{1,T}^{(1)} g \tilde{V}_T \Sigma_{TT}^{(1)} \tilde{V}_T g^{-1/2} \mathbf{1}' \Sigma_{TT}^{(1)-1}$$

For the term $\tilde{V}_T \mathcal{S}_{TT}^{(1)} \tilde{V}_T$, using (42), we have

$$\tilde{V}_T \mathcal{S}_{TT}^{(1)} \tilde{V}_T = \tilde{\#} \begin{matrix} (1)' \\ T \end{matrix} \Sigma_{TT}^{(1)-1} \begin{matrix} (1) \\ T \end{matrix} + l_1 + l_2 + l_3; \quad (44)$$

where $l_1 = \hat{\#} \begin{matrix} (1)' \\ T \end{matrix} \mathcal{S}_{TT}^{(1)-1} \begin{matrix} (1) \\ T \end{matrix}$, $l_2 = \begin{matrix} 2 \\ n \end{matrix} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right)' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right)$, $l_3 = 2 \hat{\#} \begin{matrix} (1)' \\ n \end{matrix} \begin{matrix} (1)' \\ T \end{matrix} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right)$. For the term l_1 , since j/l_1j $\hat{\#} j \begin{matrix} (1)' \\ T \end{matrix} \mathcal{S}_{TT}^{(1)-1} \begin{matrix} (1) \\ T \end{matrix}$ $\begin{matrix} (1)' \\ T \end{matrix} \Sigma_{TT}^{(1)-1} \begin{matrix} (1) \\ T \end{matrix} j + j \hat{\#} \begin{matrix} (1)' \\ T \end{matrix} \Sigma_{TT}^{(1)-1} \begin{matrix} (1) \\ T \end{matrix} j$, it follows from Lemma 4, (31), (41), and (43) that there exist constants $c_{27}, c_{28} > 0$ such that with probability at least $1 - c_{27}[(q_n S_n)^{-1} + \bar{f} \log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})]$,

$$j/l_1j \leq c_{28} \bar{f} \begin{matrix} (1)' \\ T \end{matrix} \Sigma_{TT}^{(1)-1} \begin{matrix} (1) \\ T \end{matrix} g^{-1} [q_n S_n = n + \bar{f} \log \log(n) = n g^{1/2}] + c_{28} \begin{matrix} (1)' \\ n \end{matrix} \bar{f} \begin{matrix} (1)' \\ T \end{matrix} \Sigma_{TT}^{(1)-1} \begin{matrix} (1) \\ T \end{matrix} g^{-1/2} [(q_n S_n)^{3/2} = n + \bar{f} q_n S_n \log(q_n S_n) = n g^{1/2} + \bar{f} q_n S_n \log \log(n) = n g^{1/2}]:$$

To bound the term l_2 , since $j/l_2j \lesssim \begin{matrix} 2 \\ n \end{matrix} q_n S_n \bar{f} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right)' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) g^{-1} \bar{f} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right)' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) g^{-1}$, it follows from Lemma 9 that there exist universal constants $c_{29}, c_{30} > 0$ such that with probability at least $1 - c_{29}[(q_n S_n)^{-1} + \bar{f} \log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})]$,

$$j/l_2j \leq c_{30} \begin{matrix} 2 \\ n \end{matrix} q_n S_n:$$

For the term l_3 , since $j/l_3j \leq 2 \begin{matrix} (1)' \\ n \end{matrix} \hat{\#} j \begin{matrix} (1)' \\ T \end{matrix} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) \begin{matrix} (1)' \\ T \end{matrix} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) j + 2 j \hat{\#} j \bar{f} \begin{matrix} (1)' \\ n \end{matrix} \begin{matrix} (1)' \\ T \end{matrix} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) g$, it follows from Lemma 10, (93) and (31) that there exist constants $c_{31}, c_{32} > 0$ such that with probability at least $1 - c_{31}[(q_n S_n)^{-1} + \bar{f} \log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})]$,

$$j/l_3j \leq c_{32} \left(\begin{matrix} 2 \\ n \end{matrix} q_n S_n \right)^{1/2} \bar{f} \begin{matrix} (1)' \\ T \end{matrix} \Sigma_{TT}^{(1)-1} \begin{matrix} (1) \\ T \end{matrix} g^{-1/2}:$$

By combining the above three inequalities with (44), we have that there exist universal constants $c_{33}, c_{34} > 0$ such that with probability at least $1 - c_{33}[(q_n S_n)^{-1} + \bar{f} \log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})]$,

$\exp(-n_1=12) + \exp(-n_2=12)]$,

$$j\tilde{V}_T S_{TT}^{(1)} \tilde{V}_T \leq \tilde{\#} \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} j \leq c_{34} (\frac{2}{n} q_n S_n)^{1/2} f \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} g^{-1/2}. \quad (45)$$

Since $j\tilde{V}_T S_{TT}^{(1)} \tilde{V}_T \leq \tilde{V}_T \Sigma_{TT}^{(1)} \tilde{V}_T j \leq \max(\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2}) k \Lambda_T^{(1)-1/2} (\Sigma_{TT}^{(1)} S_{TT}^{(1)}) \Lambda_T^{(1)-1/2} k_2 \tilde{V}_T S_{TT}^{(1)} \tilde{V}_T$,

it follows from Lemma 7 and Lemma 8 that there exist constants $c_{35}; c_{36} > 0$ such that

with probability at least $1 - c_{35} f(q_n S_n)^{-1} + \exp(-n_1=12) + \exp(-n_2=12)g$,

$$j\tilde{V}_T S_{TT}^{(1)} \tilde{V}_T \leq \tilde{V}_T \Sigma_{TT}^{(1)} \tilde{V}_T j \leq c_{36} f q_n^2 S_n^2 \log(q_n S_n) = n g^{1/2} \tilde{V}_T S_{TT}^{(1)} \tilde{V}_T. \quad (46)$$

For the term $\tilde{V}_T \hat{\Lambda}_T^{(1)}$, using (42) again, it has the form

$$\tilde{V}_T \hat{\Lambda}_T^{(1)} = \tilde{\#} \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} + V_1 + V_2; \quad (47)$$

where $V_1 = \hat{\#} \frac{(1)'}{T} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)} \frac{(1)}{T}$ and $V_2 = \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\frac{(1)}{T})$. Since

$jV_1 j \leq \hat{\#} j \frac{(1)'}{T} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)} \frac{(1)}{T} j + \hat{\#} j \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} j$, it follows from

Lemma 4, (31), (43) and (41) that there exist universal constants $c_{37}; c_{38} > 0$ such that

with probability at least $1 - c_{37} [(q_n S_n)^{-1} + f \log(n) g^{-1} + \exp(-n_1=12) + \exp(-n_2=12)]$,

$$jV_1 j \leq c_{38} [q_n S_n = n + f \log \log(n) = n g^{1/2}] + c_{38} \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} g^{1/2}$$

$$[(q_n S_n)^{3/2} = n + f q_n S_n \log(q_n S_n) = n g^{1/2} + f q_n S_n \log \log(n) = n g^{1/2}]:$$

Since $jV_2 j \leq \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\frac{(1)}{T}) \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\frac{(1)}{T}) j + \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\frac{(1)}{T})$,

it holds from Lemma 10, (93), and (41) that there exist constants $c_{39}; c_{40} > 0$ such that

with probability at least $1 - c_{39} [(q_n S_n)^{-1} + f \log(n) g^{-1} + \exp(-n_1=12) + \exp(-n_2=12)]$,

$$jV_2 j \leq c_{40} (\frac{2}{n} q_n S_n)^{1/2} f \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} g^{1/2}.$$

By combining the above two inequalities with (47), we conclude that there exist universal

constants $c_{41}; c_{42} > 0$ such that with probability at least $1 - c_{41} [(q_n S_n)^{-1} + f \log(n) g^{-1} +$

$\exp(-n_1=12) + \exp(-n_2=12)]$,

$$j\tilde{V}_T \hat{\Lambda}_T^{(1)} \leq \tilde{\#} \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} j \leq c_{42} (\frac{2}{n} q_n S_n)^{1/2} f \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} g^{1/2}. \quad (48)$$

Moreover, using (31), (45), (46), (48), and the fact that $\frac{2}{n}q_n S_n \frac{(1)'}{T} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} = o(1)$, elementary calculation indicates that

$$2a_n^{1/2}(U_n - b_n) = o_p(1): \quad (49)$$

For the term $\tilde{V}_T(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)})$, it follows from (42) and Holder's inequality that

$$\begin{aligned} j\tilde{V}_T(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)})j &\leq k\Lambda_T^{(1)-1/2}(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)})k_\infty q_n S_n \max(\Lambda_T^{(1)/2} S_{TT}^{(1)-1} \Lambda_T^{(1)/2})^{1/2} \\ j\hat{\#}j &\leq f_T^{(1)'} S_{TT}^{(1)-1} \hat{\cdot}_{1,T}^{(1)} g^{1/2} + \frac{1}{n} f \text{sgn}(\hat{\cdot}_{1,T}^{(1)}) \hat{\Lambda}_T^{(1)/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)/2} \text{sgn}(\hat{\cdot}_{1,T}^{(1)}) g^{1/2} : \end{aligned}$$

Together with Lemma 4, Lemma 8, Lemma 9 and (31), it can be deduced that there exist universal constants $c_{43}; c_{44} > 0$ such that with probability at least $1 - c_{43}[(q_n S_n)^{-1} + f \log(n)g^{-1} + \exp(-n_1=12) + \exp(-n_2=12)]$,

$$j\tilde{V}_T(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)})j \leq c_{44}(q_n S_n)^{1/2} f_T^{(1)'} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} g^{-1/2} k\Lambda_T^{(1)-1/2}(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)})k_\infty: \quad (50)$$

To bound the term $k\Lambda_T^{(1)-1/2}(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)})k_\infty$, note that

$$\Lambda_T^{(1)-1/2}(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)}) fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad N(0; n_1^{-1} \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}):$$

Union bound inequality and the concentration inequality imply that for any $t > 0$,

$$P \left[k\Lambda_T^{(1)-1/2}(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)})k_\infty \geq t fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \right] \leq 2q_n S_n \exp(-\frac{1}{4}nt^2g):$$

Plugging $t = c_{45} f \log(q_n S_n) = ng^{1/2}$ with $c_{45} = (8 - \frac{1}{2})^{1/2}$ into the above yields $P \left[k\Lambda_T^{(1)-1/2}(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)})k_\infty \geq c_{45} f \log(q_n S_n) = ng^{1/2} fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \right] \leq 2(q_n S_n)^{-1}$. Together with

Lemma 3, it can be deduced that $P \left[k\Lambda_T^{(1)-1/2}(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)})k_\infty \geq c_{45} f \log(q_n S_n) = ng^{1/2}$

$1 - 2f(q_n S_n)^{-1} + \exp(-n_1=12) + \exp(-n_2=12)g$. Together with (50), there exist universal constants $c_{46}; c_{47} > 0$ such that with probability at least $1 - c_{46}[(q_n S_n)^{-1} + f \log(n)g^{-1} + \exp(-n_1=12) + \exp(-n_2=12)]$,

$$j\tilde{V}_T(\hat{\cdot}_{1,T}^{(1)} \quad \hat{\cdot}_{1,T}^{(1)})j \leq c_{47} f_T^{(1)'} \Sigma_{TT}^{(1)-1} \frac{(1)}{T} g^{-1/2} f q_n S_n \log(q_n S_n) = ng^{1/2}:$$

Together with (31), (45), (46), (48), and conditions (C2)–(C5), it is seen that $4a_n X_n = o_p(1)$. Together with (49), (41), (40), and Lemma 12, it can be concluded that

$$\Omega_1 = \Omega_1^* \stackrel{P}{\rightarrow} 1; \quad \Omega_1^* \neq 0; \quad (51)$$

Similar argument leads to $\Omega_2 = \Omega_2^* \stackrel{P}{\rightarrow} 1$, $\Omega_2^* \neq 0$. Together with (38), (39), and (51), it holds that $R^\circ(\hat{v}) = R^\circ(\hat{v}^{(1)}) \stackrel{P}{\rightarrow} 1$, $R^\circ(\hat{v}^{(1)}) \neq 0$, which completes the proof. \square

Proof of Corollary 1: It follows directly from Theorems 1 and 2. \square

In the next section, we present all the auxiliary lemmas with their proofs.

3 Auxiliary lemmas with their proofs

Lemma 1. *Assume the following conditions (a)–(b):*

$$(a) \quad c_1 = \min(\Lambda^{\dagger 1/2} \Sigma \Lambda^{\dagger 1/2}) \quad \max(\Lambda^{\dagger 1/2} \Sigma \Lambda^{\dagger 1/2}) \leq c_2,$$

$$c_1 = \min(\Lambda^{(1)-1/2} \Sigma^{(1)} \Lambda^{(1)-1/2}) \quad \max(\Lambda^{(1)-1/2} \Sigma^{(1)} \Lambda^{(1)-1/2}) \leq c_2,$$

for some universal constants $0 < c_1 < c_2$.

$$(b) \quad \mathbb{P}_{j \in T^*} \mathbb{P}_{k=s_n+1}^\infty \sum_{jk}^* = o(\min_{j \in T^*} \sum_{k=1}^{s_n} \sum_{jk}^*).$$

Then we have the following properties:

$$1) \quad N = N^* \text{ and } T^* = T.$$

$$2) \quad \Delta^{(1)2} = f_1 + o(r_n^{-1}) + o(r_n^{-1/2} \frac{1}{n}) g \Delta^2,$$

where the parameter $\eta_n = \left(\begin{matrix} *^{(1)'} & \Sigma_{TT}^{(1,2)} & \Sigma_{TT}^{(2)\dagger} & \Sigma_{TT}^{(2,1)} \\ T & & & T \end{matrix} \right) = \left(\begin{matrix} *^{(1)'} & \Sigma_{TT}^{(1)} \\ T & T \end{matrix} \right) \leq 1$.

Together with the triangle inequality, we have

$$\begin{aligned}
& k\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \quad {}^*(1)k_2 \\
& k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)} \quad {}^*(2)k_2 + k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1,2)} \quad {}^*(2)k_2 + \\
& k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)} \quad {}^*(2)k_2 + k\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)} \quad {}^*(2)k_2;
\end{aligned}$$

which further implies that

$$\begin{aligned}
& k\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \quad {}^*(1)k_2^2 \\
& \lesssim k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TT}^{(1,2)} \quad {}^*(2)k_2^2 + k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1,2)} \quad {}^*(2)k_2^2 + \\
& k\Lambda_N^{(1)-1/2}\Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1,2)} \quad {}^*(2)k_2^2 + k\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)} \quad {}^*(2)k_2^2. \tag{54}
\end{aligned}$$

Based on condition (a) and Lemma 14, it is trivial to show that

$$\begin{aligned}
& \min \Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2} \\
& = \frac{-1}{\max} \Lambda_N^{(1)1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})^{-1}\Lambda_N^{(1)1/2} \quad c_1; \tag{55}
\end{aligned}$$

for the universal constant $c_1 > 0$ defined in condition (a). Hence, for the term $k\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \quad {}^*(1)k_2^2$, we have

$$\begin{aligned}
& k\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \quad {}^*(1)k_2^2 \\
& c_1 \max \Lambda_N^{(1)1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})^{-1}\Lambda_N^{(1)1/2} \quad f(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \quad {}^*(1)g \\
& \quad \Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2}f(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)}) \quad {}^*(1)g \\
& c_1(\Lambda_N^{(1)1/2} \quad {}^*(1))'f\Lambda_N^{(1)-1/2}(\Sigma_{NN}^{(1)} \quad \Sigma_{NT}^{(1)}\Sigma_{TT}^{(1)-1}\Sigma_{TN}^{(1)})\Lambda_N^{(1)-1/2}g(\Lambda_N^{(1)1/2} \quad {}^*(1)) \\
& c_1^2 k\Lambda_N^{(1)1/2} \quad {}^*(1)k_2^2. \tag{56}
\end{aligned}$$

where the first inequality is by (55), and the last inequality is also based on (55). According

to condition (a) and Lemma 14 again, we have

$$\min \Lambda_N^{(1)1/2} \Sigma_{NN}^{(1)-1} \Lambda_N^{(1)1/2} = \max \Lambda_N^{(1)-1/2} \Sigma_{NN}^{(1)} \Lambda_N^{(1)-1/2} \quad \mathcal{C}_2^{-1}, \quad (57)$$

$$\min \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} \quad \mathcal{C}_1, \quad (58)$$

$$\max \Lambda_T^{(1)-1/2} \Sigma_{TN}^{(1)} \Sigma_{NN}^{(1)-1} \Sigma_{NT}^{(1)} \Lambda_T^{(1)-1/2} \quad \max \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} \quad \mathcal{C}_2, \quad (59)$$

$$\max \Lambda_T^{(2)\dagger 1/2} \Sigma_{TT}^{(2,1)} \Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Lambda_T^{(2)\dagger 1/2} \quad \max \Lambda_T^{(2)\dagger 1/2} \Sigma_{TT}^{(2)} \Lambda_T^{(2)\dagger 1/2} \quad \mathcal{C}_2, \quad (60)$$

for the universal constants \mathcal{C}_1 and \mathcal{C}_2 defined in condition (a). Thus, for the term

$k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)} k_2^2$, we have

$$\begin{aligned} & k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)} k_2^2 \\ & \mathcal{C}_2 \min \Lambda_N^{(1)1/2} \Sigma_{NN}^{(1)-1} \Lambda_N^{(1)1/2} (\Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)})' (\Lambda_N^{(1)-1/2} \Lambda_N^{(1)-1/2}) (\Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)}) \\ & \mathcal{C}_2 (\Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)})' \Sigma_{NN}^{(1)-1} (\Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)}) \\ & \mathcal{C}_2 \max \Lambda_T^{(1)-1/2} \Sigma_{TN}^{(1)} \Sigma_{NN}^{(1)-1} \Sigma_{NT}^{(1)} \Lambda_T^{(1)-1/2} \quad k\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)} k_2^2 \\ & \mathcal{C}_2^2 (\Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)})' (\Lambda_T^{(1)1/2} \Lambda_T^{(1)1/2}) (\Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)}) \\ & \mathcal{C}_2^2 \mathcal{C}_1^{-1} \min \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} (\Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)})' (\Lambda_T^{(1)1/2} \Lambda_T^{(1)1/2}) (\Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Sigma_T^{*(2)}) \\ & \mathcal{C}_2^2 \mathcal{C}_1^{-1} \max \Lambda_T^{(2)\dagger 1/2} \Sigma_{TT}^{(2,1)} \Sigma_{TT}^{(1)-1} \Sigma_{TT}^{(1,2)} \Lambda_T^{(2)\dagger 1/2} \quad k\Lambda_T^{(2)1/2} \Sigma_T^{*(2)} k_2^2 \\ & \mathcal{C}_2^3 \mathcal{C}_1^{-1} k\Lambda_T^{(2)1/2} \Sigma_T^{*(2)} k_2^2, \end{aligned}$$

where the first inequality is by (57), the fourth inequality follows from (59), the fifth inequality is based on (58), and the last inequality is according to (60). Likewise, for the term $k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1,2)} \Sigma_T^{*(2)} k_2^2$, we have

$$k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1,2)} \Sigma_T^{*(2)} k_2^2 \quad \mathcal{C}_2^2 k\Lambda_T^{(2)1/2} \Sigma_T^{*(2)} k_2^2.$$

In a similar fashion, for the term $k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TN}^{(1,2)} \Sigma_N^{*(2)} k_2^2$, we have

$$k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Sigma_{TN}^{(1,2)} \Sigma_N^{*(2)} k_2^2 \quad \mathcal{C}_2^3 \mathcal{C}_1^{-1} k\Lambda_N^{(2)1/2} \Sigma_N^{*(2)} k_2^2.$$

In addition, for the term $k\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)}{}^*(2)k_2^2$, one has

$$k\Lambda_N^{(1)-1/2}\Sigma_{NN}^{(1,2)}{}^*(2)k_2^2 = (\Sigma_{NN}^{(1,2)}{}^*(2))'(\Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2})(\Sigma_{NN}^{(1,2)}{}^*(2))$$

$$\mathcal{C}_2 \min \Lambda_N^{(1)1/2}\Sigma_{NN}^{(1)-1}\Lambda_N^{(1)1/2}(\Sigma_{NN}^{(1,2)}{}^*(2))'(\Lambda_N^{(1)-1/2}\Lambda_N^{(1)-1/2})(\Sigma_{NN}^{(1,2)}{}^*(2))$$

$$\mathcal{C}_2(\Lambda_N^{(2)1/2}{}^*(2))'(\Lambda_N^{(2)\dagger 1/2}\Sigma_{NN}^{(2,1)}\Sigma_{NN}^{(1)-1}\Sigma_{NN}^{(1,2)}\Lambda_N^{(2)\dagger 1/2})(\Lambda_N^{(2)1/2}{}^*(2))$$

$$\mathcal{C}_2 \max \Lambda_N^{(2)\dagger 1/2}\Sigma_{NN}^{(2,1)}\Sigma_{NN}^{(1)-1}\Sigma_{NN}^{(1,2)}\Lambda_N^{(2)\dagger 1/2} k\Lambda_N^{(2)1/2}{}^*(2)k_2^2$$

$$\mathcal{C}_2 \max \Lambda_N^{(2)\dagger 1/2}\Sigma_{NN}^{(2)}\Lambda_N^{(2)\dagger 1/2} k\Lambda_N^{(2)1/2}{}^*(2)k_2^2$$

\mathcal{C}

N

where the second equality follows from (61), and the last equality is based on (63) and (64). Together with (64), (63) and (62), it can be concluded that $\Delta^{(1)2} = \bar{f}1 + o(r_n^{-1}) + o(r_n^{-1/2} \frac{1}{n} g \Delta^2)$, which completes the proof. \square

Lemma 2. Assume the invertibility of $S_{TT}^{(1)}$ and consider the following optimization problem:

$$\min_{v_T \in \mathbb{R}^{qn^{sn}}} \frac{1}{2} v_T' S_{TT}^{(1)} v_T + \frac{n_1 n_2}{n(n-2)} \hat{\Lambda}_T^{(1)1/2} v_T' \hat{\Lambda}_T^{(1)1/2} v_T + \frac{n_1 n_2}{n(n-2)} v_T' \hat{\Lambda}_T^{(1)1/2} v_T \text{sgn}(\hat{\Lambda}_T^{(1)1/2} v_T) ;$$

where $v_T = (v_1' \dots v_{qn}')'$ with sub-vectors $v_j = (v_{j1} \dots v_{jsn})' \in \mathbb{R}^{sn}$. Let \tilde{v}_T be the solution of this optimization problem where $\tilde{v}_T = (\tilde{v}_1' \dots \tilde{v}_{qn}')'$ with sub-vectors $\tilde{v}_j = (\tilde{v}_{j1} \dots \tilde{v}_{jsn})' \in \mathbb{R}^{sn}$, then we have:

$$\begin{aligned} \tilde{v}_T &= \bar{f} n_1 n_2 n^{-1} (n-2)^{-1} g \bar{f} 1 + \frac{n_1 n_2}{n(n-2)} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\hat{\Lambda}_T^{(1)1/2} v_T) g \\ &= \bar{f} n_1 n_2 n^{-1} (n-2)^{-1} g \hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\hat{\Lambda}_T^{(1)1/2} v_T) : \end{aligned}$$

Proof of Lemma 2: The proof is analogous to that of Lemma 16. \square

Lemma 3. Define the events M_n and M_n^* as

$$\begin{aligned} M_n &= \{ \bar{f}_1 = 2, n_1 = n-3, \bar{f}_2 = 2, n_2 = n-3 \} \\ M_n^* &= \{ \bar{f}_1 = 2, n_1 n_2 = n^2, \bar{f}_2 = 4 \} \end{aligned}$$

Then we have the following properties:

- 1) $P(M_n) \leq 2 \exp(-n-12) + 2 \exp(-n-2)$.
- 2) $P(M_n^*) \leq 2 \exp(-n-12) + 2 \exp(-n-2)$.

Proof of Lemma 3: First of all, note that $n_1 \sim \text{Binomial}(n; \frac{1}{2})$. Invoking the Chernoff tail bounds for binomial random variables, we have that for any $\epsilon \in [0; 1]$,

$$\begin{aligned} P(\bar{f} n_1 \leq (1-\epsilon)n) &\leq \exp(-n\epsilon^2/3); \\ P(\bar{f} n_1 \geq (1+\epsilon)n) &\leq \exp(-n\epsilon^2/3); \end{aligned}$$

Then, we substitute $\epsilon = 1/2$ into the above two inequalities to obtain

$$\begin{aligned} P(n_1 = n - 3 \mid \epsilon = 1/2) &\leq \exp(-n/12); \\ P(n_1 = n - 1 \mid \epsilon = 1/2) &\leq \exp(-n/12); \end{aligned} \tag{65}$$

Accordingly, we have

$$\begin{aligned} P(\epsilon = 1/2 \mid n_1 = n - 3) &\leq P(n_1 = n - 3 \mid \epsilon = 1/2) \\ &\leq \exp(-n/12); \end{aligned}$$

where the last inequality is by (65). By symmetry, one has

$$P(\epsilon = 1/2 \mid n_2 = n - 3) \leq \exp(-n/12);$$

To this end, based on the above two inequalities, we can deduce that $P(\mathcal{M}_n) \leq \exp(-n/12) + \exp(-n/12)$, which completes the proof of 1). Property 2) follows from the fact that $\mathcal{M}_n \subseteq \mathcal{M}_n^*$. \square

Lemma 4. For any $\epsilon \in (e^{-n/100}, 1/100)$, define the event $\mathcal{M}_{3n}(\epsilon)$ as

$$\begin{aligned} \mathcal{M}_{3n}(\epsilon) = & \bigcap_{T \in \mathcal{T}} \left\{ \sum_{T'} S_{TT'}^{(1)-1} \leq n + \log(\epsilon^{-1}) \right\} \\ & + \bigcap_{T \in \mathcal{T}} \left\{ \sum_{T'} S_{TT'}^{(1)-1} \leq n + \log(\epsilon^{-1}) \right\} \end{aligned}$$

Proof of Lemma 4: First of all, note that

$$\begin{aligned}
& \hat{\Sigma}_T^{(1)'} \mathcal{S}_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \\
= & (\hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)}) (\hat{\Sigma}_T^{(1)'} \mathcal{S}_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} = \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad 1) \\
& + \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} (\hat{\Sigma}_T^{(1)'} \mathcal{S}_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} = \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad 1) + (\hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)});
\end{aligned}$$

which implies that

$$\begin{aligned}
& j \hat{\Sigma}_T^{(1)'} \mathcal{S}_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} j \\
& j \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} j \quad j \hat{\Sigma}_T^{(1)'} \mathcal{S}_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} = \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad 1j \\
& + \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} j \hat{\Sigma}_T^{(1)'} \mathcal{S}_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} = \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad 1j + j \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} j;
\end{aligned}$$

Together with Lemma 18 and Lemma 19, we conclude that with probability at least 1

$$1 - 4 \exp(-n_1/12) - 4 \exp(-n_2/12),$$

$$\begin{aligned}
j \hat{\Sigma}_T^{(1)'} \mathcal{S}_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} \quad \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} j & \lesssim q_n S_n = n + \log(\%^{-1}) = n + \quad q_n S_n = n + f \log(\%^{-1}) = ng^{1/2} \\
& f \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} g + f \log(\%^{-1}) = ng^{1/2} f \hat{\Sigma}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T^{(1)} g^{1/2};
\end{aligned}$$

which completes the proof. \square

Lemma 5. *Assume the following condition (a):*

$$(a) \log(q_n S_n) = o(n).$$

Then there exist universal constants $c_1 > 0$ and $c_2 > 0$ such that:

$$\begin{aligned}
1) P \left(k \hat{\Lambda}_T^{(1)} \hat{\Lambda}_T^{(1)-1} \quad l_{q_n S_n} k_{\max} \quad c_1 f \log(q_n S_n) = ng^{1/2} \quad 1 \quad c_2 f(q_n S_n)^{-1} \right. \\
\left. + \exp(-n_1/12) + \exp(-n_2/12) \right) g.
\end{aligned}$$

$$\begin{aligned}
2) P \left(k \hat{\Lambda}_T^{(1)} \hat{\Lambda}_T^{(1)-1} \quad l_{q_n S_n} k_{\max} \quad c_1 f \log(q_n S_n) = ng^{1/2} \quad 1 \quad c_2 f(q_n S_n)^{-1} \right. \\
\left. + \exp(-n_1/12) + \exp(-n_2/12) \right) g.
\end{aligned}$$

$$3) P k \hat{\Lambda}_T^{(1)1/2} \Lambda_T^{(1)-1/2} I_{q_n s_n} k_{\max} c_1 f \log(q_n s_n) = n g^{1/2} 1 c_2 f(q_n s_n)^{-1} \\ + \exp(n_1=12) + \exp(n_2=12)g.$$

$$4) P k \Lambda_T^{(1)1/2} \hat{\Lambda}_T^{(1)-1/2} I_{q_n s_n} k_{\max} c_1 f \log(q_n s_n) = n g^{1/2} 1 c_2 f(q_n s_n)^{-1} \\ + \exp(n_1=12) + \exp(n_2=12)g.$$

Note that $I_{q_n s_n}$ denotes the $q_n s_n \times q_n s_n$ identity matrix.

Proof of Lemma 5: Before showing the Lemma, we prepare some notations. For any sub-exponential random variable X , its sub-exponential norm is denoted as $kXk_\psi = \sup_{q \geq 1} q^{-1} fE(jXj^q)g^{1/q}$. Now, we are in a position to start the proof. First of all, notice that

$$k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} I_{q_n s_n} k_{\max} = \max_{j \in T} \max_{k \leq s_n} j!_{jk}!_{jk}^{-1} 1j; \quad (66)$$

Moreover, by definition, we have that for every $j \in T$ and $k \in S_n$,

$$j!_{jk} = (n_2)^{-1} n_1 \prod_{i \in H_1} (i_{jk} - 1_{jk})^2 = n_1 + n_2 \prod_{i' \in H_2} (i'_{jk} - 2_{jk})^2 = n_2 \\ (n_2)^{-1} n_1 \prod_{i_1 \in H_1} (i_1 j k = n_1 - 1_{jk})^2 + n_2 \prod_{i_2 \in H_2} (i_2 j k = n_2 - 2_{jk})^2 ;$$

which implies that for every $j \in T$ and $k \in S_n$,

$$j!_{jk}!_{jk}^{-1} 1 = (n_2)^{-1} n_1 n_1^{-1} \prod_{i \in H_1} f!_{jk}^{-1/2} (i_{jk} - 1_{jk}) g^2 1 \\ + (n_2)^{-1} n_2 n_2^{-1} \prod_{i' \in H_2} f!_{jk}^{-1/2} (i'_{jk} - 2_{jk}) g^2 1 \\ (n_2)^{-1} n_1 n_1^{-1} \prod_{i_1 \in H_1} !_{jk}^{-1/2} (i_1 j k - 1_{jk})^2 \\ (n_2)^{-1} n_2 n_2^{-1} \prod_{i_2 \in H_2} !_{jk}^{-1/2} (i_2 j k - 2_{jk})^2 + 2(n_2)^{-1};$$

Together with (66), we obtain

$$k\hat{\Lambda}_T^{(1)}\Lambda_T^{(1)-1} \leq l_{qns_n}k_{\max} \leq 2n^{-1}n_1\Upsilon_1 + 2n^{-1}n_2\Upsilon_2 + 2n^{-1}n_1\Upsilon_3^2 + 2n^{-1}n_2\Upsilon_4^2 + 3n^{-1}; \quad (67)$$

where

$$\begin{aligned} \Upsilon_1 &= \max_{j \in T} \max_{k \leq s_n} n_1^{-1} \times_{i \in H_1} f!_{jk}^{-1/2}(\mathbf{ijk} \quad \mathbf{1jk}) \mathcal{G}^2 \quad 1; \\ \Upsilon_2 &= \max_{j \in T} \max_{k \leq s_n} n_2^{-1} \times_{i' \in H_2} f!_{jk}^{-1/2}(\mathbf{i'jk} \quad \mathbf{2jk}) \mathcal{G}^2 \quad 1; \\ \Upsilon_3 &= \max_{j \in T} \max_{k \leq s_n} n_1^{-1} \times_{i_1 \in H_1} !_{jk}^{-1/2}(\mathbf{i_1jk} \quad \mathbf{1jk}); \\ \Upsilon_4 &= \max_{j \in T} \max_{k \leq s_n} n_2^{-1} \times_{i_2 \in H_2} !_{jk}^{-1/2}(\mathbf{i_2jk} \quad \mathbf{2jk}); \end{aligned}$$

At this point, note that for every $i \geq H_1; j \leq q_n; k \leq s_n$, the sub-exponential norms of the sub-exponential random variables $f!_{jk}^{-1/2}(\mathbf{ijk} \quad \mathbf{1jk}) \mathcal{G}^2$ satisfy

$$k f!_{jk}^{-1/2}(\mathbf{ijk} \quad \mathbf{1jk}) \mathcal{G}^2 k_\psi \leq \max f!_{jk}^{-1/2}(\mathbf{ijk} \quad \mathbf{1jk}) \mathcal{G}^2 \leq 2e^{2/e} g. \quad (68)$$

For the term Υ_1 , conditional on any nonempty $fY_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n$, one can show that for any $t > 0$,

$$\begin{aligned} P \Upsilon_1 \leq t \quad & P \left(\times_{j \in T} \times_{k \leq s_n} n_1^{-1} \times_{i \in H_1} f!_{jk}^{-1/2}(\mathbf{ijk} \quad \mathbf{1jk}) \mathcal{G}^2 \leq 1 \mid t fY_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \right) \\ & \leq 2q_n s_n \exp(-c_1 \min f!_{jk}^{-1/2}(\mathbf{ijk} \quad \mathbf{1jk}) \mathcal{G}^2; t g n); \end{aligned} \quad (69)$$

for some universal constant $c_1 > 0$, where the first inequality holds from the union bound inequality, and the second inequality follows from (68) and the Bernstein inequality in Lemma H.2 of Ning and Liu (2017). Similar reasoning gives the result that for any $t > 0$,

$$P \Upsilon_2 \leq t \quad P \left(\times_{j \in T} \times_{k \leq s_n} n_2^{-1} \times_{i' \in H_2} f!_{jk}^{-1/2}(\mathbf{i'jk} \quad \mathbf{2jk}) \mathcal{G}^2 \leq 1 \mid t fY_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \right) \leq 2q_n s_n \exp(-c_2 \min f!_{jk}^{-1/2}(\mathbf{i'jk} \quad \mathbf{2jk}) \mathcal{G}^2; t g n); \quad (70)$$

for some universal constant $c_2 > 0$. Regarding the term Υ_3 , it is clear that for any $t > 0$,

$$\begin{aligned}
& P \left\{ \Upsilon_3 \geq t \mid \mathcal{F}_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \right\} \\
& \leq \sum_{j \in T} \sum_{k \leq s_n} \sum_{i_1 \in H_1} \sum_{i_2 \in H_2} \dots \sum_{i_k \in H_k} \prod_{j \in T} P \left\{ n_1^{-1} \sum_{i_1 \in H_1} \dots \sum_{i_k \in H_k} \left| \sum_{j \in T} \left(\sum_{i_1 \in H_1} \dots \sum_{i_k \in H_k} \right) \right| \right\} \\
& \leq 2q_n s_n \exp(-c_3 n t^2); \tag{71}
\end{aligned}$$

for some universal constant $c_3 > 0$, where the first inequality is based on the union bound inequality, and the second inequality follows from Hoeffding inequality. Similar argument leads to the result that for any $t > 0$,

$$P \left\{ \Upsilon_4 \geq t \mid \mathcal{F}_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \right\} \leq 2q_n s_n \exp(-c_4 n t^2); \tag{72}$$

for some universal constant $c_4 > 0$. To this end, conditional on any nonempty $\mathcal{F}_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n$, it can be deduced that for any $t > 0$,

$$\begin{aligned}
& P \left\{ k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} \geq I_{q_n s_n} k_{\max} \mid \mathcal{F}_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \right\} \\
& \leq P \left\{ 2n^{-1} n_1 \Upsilon_1 + 2n^{-1} n_2 \Upsilon_2 + 2n^{-1} n_1 \Upsilon_3^2 + 2n^{-1} n_2 \Upsilon_4^2 + 3n^{-1} \sum_{i \in T} \left| \sum_{j \in T} \left(\sum_{i_1 \in H_1} \dots \sum_{i_k \in H_k} \right) \right| \right\} \\
& \leq P \left\{ \Upsilon_1 + \Upsilon_2 + \Upsilon_3^2 + \Upsilon_4^2 + n^{-1} \sum_{i \in T} \left| \sum_{j \in T} \left(\sum_{i_1 \in H_1} \dots \sum_{i_k \in H_k} \right) \right| \right\} \\
& \leq P \left\{ \Upsilon_1 \geq 5^{-1} c_5 t \mid \mathcal{F}_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \right\} + P \left\{ \Upsilon_2 \geq 5^{-1} c_5 t \mid \mathcal{F}_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \right\} \\
& \quad + P \left\{ \Upsilon_3 \geq 5^{-1/2} c_5^{1/2} t^{1/2} \mid \mathcal{F}_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \right\} \\
& \quad + P \left\{ \Upsilon_4 \geq 5^{-1/2} c_5^{1/2} t^{1/2} \mid \mathcal{F}_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \right\} \\
& \quad + P \left\{ n^{-1} \sum_{i \in T} \left| \sum_{j \in T} \left(\sum_{i_1 \in H_1} \dots \sum_{i_k \in H_k} \right) \right| \geq 5^{-1} c_5 t \mid \mathcal{F}_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \right\} \\
& \leq 4q_n s_n \exp(-c_6 \min\{t^2, t\}) + 4q_n s_n \exp(-c_6 n t) + P(n^{-1} \sum_{i \in T} \left| \sum_{j \in T} \left(\sum_{i_1 \in H_1} \dots \sum_{i_k \in H_k} \right) \right| \geq 5^{-1} c_5 t) \\
& \leq 8q_n s_n \exp(-c_6 \min\{t^2, t\}) + P(n^{-1} \sum_{i \in T} \left| \sum_{j \in T} \left(\sum_{i_1 \in H_1} \dots \sum_{i_k \in H_k} \right) \right| \geq 5^{-1} c_5 t);
\end{aligned}$$

for some carefully chosen universal constants $c_5 > 0$ and $c_6 > 0$, where the first inequality is by (67), the second inequality comes from the definition of \mathcal{M}_n in Lemma 3, the fourth

inequality is based on (69), (70), (71) and (72). Accordingly, we set $c_7 = (2c_6^{-1})^{1/2}$ and substitute $t = c_7 \mathbb{1} \log(q_n s_n) = n g^{1/2}$ into the above inequality to obtain

$$P \left(k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} \leq I_{q_n s_n} k_{\max} \quad c_7 \mathbb{1} \log(q_n s_n) = n g^{1/2} \quad \mathbb{1} \quad \mathbb{1} \quad 8(q_n s_n)^{-1} : \right. \\ \left. \right) \quad (73)$$

It then follows that

$$P \left(k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} \leq I_{q_n s_n} k_{\max} \quad c_7 \mathbb{1} \log(q_n s_n) = n g^{1/2} \right. \\ \times \\ \left. P \left(k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} \leq I_{q_n s_n} k_{\max} \quad c_7 \mathbb{1} \log(q_n s_n) = n g^{1/2} \quad \mathbb{1} \quad \mathbb{1} \quad 8(q_n s_n)^{-1} : \right. \right. \\ \left. \left. \{y_i\}_{i=1}^n \in \mathcal{M}_n \right) \right. \\ \times \\ \left. P \left(\mathbb{1} \quad 8(q_n s_n)^{-1} g \quad \mathbb{1} \quad \mathbb{1} \quad 8(q_n s_n)^{-1} g P(\mathcal{M}_n) \right. \right. \\ \left. \left. \{y_i\}_{i=1}^n \in \mathcal{M}_n \right) \right. \\ \left. \mathbb{1} \quad 8 \mathbb{1} \quad 8(q_n s_n)^{-1} + \exp(-n) + \exp(-n) \right) g;$$

where the second inequality is by (73), and the last inequality follows from Lemma 3.

Therefore, property 1) holds from the above inequality. Moreover, it can be verified that

under the event $k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} \leq I_{q_n s_n} k_{\max} \quad c_7 \mathbb{1} \log(q_n s_n) = n g^{1/2}$,

$$k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} \leq I_{q_n s_n} k_{\max} \quad 2k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} \leq I_{q_n s_n} k_{\max}:$$

Hence, based on the above two inequalities, we conclude that

$$P \left(k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} \leq I_{q_n s_n} k_{\max} \quad 2c_7 \mathbb{1} \log(q_n s_n) = n g^{1/2} \right. \\ \left. \mathbb{1} \quad 8 \mathbb{1} \quad 8(q_n s_n)^{-1} + \exp(-n) + \exp(-n) \right) g; \quad (74)$$

which completes the proof of property 2). Property 3) can be directly proved by using the fact that $k \hat{\Lambda}_T^{(1)1/2} \Lambda_T^{(1)-1/2} \leq I_{q_n s_n} k_{\max} \quad k \hat{\Lambda}_T^{(1)} \Lambda_T^{(1)-1} \leq I_{q_n s_n} k_{\max}$. Likewise, one can show property 4), which finishes the proof. \square

Lemma 6. Assume the following conditions (a)-(b):

$$(a) \sup_{j \leq p_n} \prod_{k=1}^{\infty} I_{jk} < 1, \quad \min(\Lambda_N^{(1)}) \geq c_0 S_n^{-a} \text{ for some constants } c_0 > 0 \text{ and } a > 1.$$

$$(b) S_n^{2a} \log f(\rho_n, q_n) S_n \mathcal{G} = o(n).$$

Then there exist universal constants $c_1 > 0$ and $c_2 > 0$ such that:

$$1) P \left[k \hat{\Lambda}_N^{(1)} \Lambda_N^{(1)-1} I_{(p_n - q_n) S_n} k_{\max} \geq c_1 [\log f(\rho_n, q_n) S_n \mathcal{G} = n]^{1/2} \right] \leq c_2 [f(\rho_n, q_n) S_n \mathcal{G}^{-1} + \exp(-n_1/12) + \exp(-n_2/12)].$$

$$2) P \left[k \Lambda_N^{(1)} \hat{\Lambda}_N^{(1)-1} I_{(p_n - q_n) S_n} k_{\max} \geq c_1 [\log f(\rho_n, q_n) S_n \mathcal{G} = n]^{1/2} \right] \leq c_2 [f(\rho_n, q_n) S_n \mathcal{G}^{-1} + \exp(-n_1/12) + \exp(-n_2/12)].$$

$$3) P \left[k \hat{\Lambda}_N^{(1)1/2} \Lambda_N^{(1)-1/2} I_{(p_n - q_n) S_n} k_{\max} \geq c_1 [\log f(\rho_n, q_n) S_n \mathcal{G} = n]^{1/2} \right] \leq c_2 [f(\rho_n, q_n) S_n \mathcal{G}^{-1} + \exp(-n_1/12) + \exp(-n_2/12)].$$

$$4) P \left[k \Lambda_N^{(1)1/2} \hat{\Lambda}_N^{(1)-1/2} I_{(p_n - q_n) S_n} k_{\max} \geq c_1 [\log f(\rho_n, q_n) S_n \mathcal{G} = n]^{1/2} \right] \leq c_2 [f(\rho_n, q_n) S_n \mathcal{G}^{-1} + \exp(-n_1/12) + \exp(-n_2/12)].$$

$$5) P \left[\det(\hat{\Lambda}_N^{(1)}) \notin 0 \right] \leq c_2 [f(\rho_n, q_n) S_n \mathcal{G}^{-1} + \exp(-n_1/12) + \exp(-n_2/12)].$$

Note that $I_{(p_n - q_n) S_n}$ denotes the $(\rho_n, q_n) S_n \times (\rho_n, q_n) S_n$ identity matrix.

Proof of Lemma 6: The proof of property 1) is analogous to that of property 1) in Lemma 5.

Then, it can be deduced that there exists $c_3 > 0$ and $c_4 > 0$ such that with probability at

least $1 - c_3 [f(\rho_n, q_n) S_n \mathcal{G}^{-1} + \exp(-n_1/12) + \exp(-n_2/12)]$,

$$\min(\hat{\Lambda}_N^{(1)}) \geq \min(\Lambda_N^{(1)}) \geq \max(\Lambda_N^{(1)}) k \hat{\Lambda}_N^{(1)} \Lambda_N^{(1)-1} I_{(p_n - q_n) S_n} k_{\max} \geq c_4 S_n^{-a};$$

where the last inequality is based on (a), (b) and property 1). As a result, property 5) holds true from the above inequality. Finally, properties 2) to 4) can be derived in a similar fashion as properties 2) to 4) in Lemma 5, which finishes the proof. \square

Lemma 7. *Assume the following condition (a):*

$$(a) \log(q_n S_n) = o(n).$$

Then there exist universal constants $c_1 > 0$ and $c_2 > 0$ such that:

$$P \leq c_1 \Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} k_2 \log(q_n S_n) = n g^{1/2} \\ + c_2 f(q_n S_n)^{-1} + \exp(-n_1/12) + \exp(-n_2/12) g;$$

Proof of Lemma 7: First of all, we note that

$$k \Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} k_2 = \Omega_1 + \Omega_2 + \Omega_3 + \Omega_4 + \Omega_5; \quad (75)$$

where

$$\begin{aligned} \Omega_1 &= 2n^{-1} n_1 q_n S_n \max_{j_2 \in T} \max_{k_2 \leq s_n} \max_{j_1 \in T} \max_{k_1 \leq s_n} n_1^{-1} \times \prod_{i \in H_1} h_{j_1 k_1}^{-1/2} (i j_1 k_1 \quad 1 j_1 k_1) \\ &\quad \times \prod_{i \in H_1} h_{j_2 k_2}^{-1/2} (i j_2 k_2 \quad 1 j_2 k_2) \times \text{corr}(j_1 k_1; j_2 k_2); \\ \Omega_2 &= 2n^{-1} n_2 q_n S_n \max_{j_2 \in T} \max_{k_2 \leq s_n} \max_{j_1 \in T} \max_{k_1 \leq s_n} n_2^{-1} \times \prod_{i \in H_2} h_{j_1 k_1}^{-1/2} (i j_1 k_1 \quad 2 j_1 k_1) \\ &\quad \times \prod_{i \in H_2} h_{j_2 k_2}^{-1/2} (i j_2 k_2 \quad 2 j_2 k_2) \times \text{corr}(j_1 k_1; j_2 k_2); \\ \Omega_3 &= 2n^{-1} n_1 q_n S_n \max_{j_2 \in T} \max_{k_2 \leq s_n} \max_{j_1 \in T} \max_{k_1 \leq s_n} n_1^{-1} \times \prod_{i_1 \in H_1} h_{i_1 j_1 k_1}^{-1/2} (i_1 j_1 k_1 \quad 1 j_1 k_1) \\ &\quad \times \prod_{i_1 \in H_1} h_{j_2 k_2}^{-1/2} (i_1 j_2 k_2 \quad 1 j_2 k_2); \\ \Omega_4 &= 2n^{-1} n_2 q_n S_n \max_{j_2 \in T} \max_{k_2 \leq s_n} \max_{j_1 \in T} \max_{k_1 \leq s_n} n_2^{-1} \times \prod_{i_2 \in H_2} h_{i_2 j_1 k_1}^{-1/2} (i_2 j_1 k_1 \quad 2 j_1 k_1) \\ &\quad \times \prod_{i_2 \in H_2} h_{j_2 k_2}^{-1/2} (i_2 j_2 k_2 \quad 2 j_2 k_2); \\ \Omega_5 &= 4n^{-1} q_n S_n; \end{aligned}$$

$t = 0$,

$$\begin{aligned}
& P \Omega_3 \quad t fY_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \\
& \times \prod_{j_2 \in T} \prod_{k_2=1} \prod_{j_1 \in T} \prod_{k_1=1} \prod_{i_1 \in H_1} P \quad n_1^{-1} \times \prod_{j_1 k_1}^{-1/2} (i_1 j_1 k_1 \quad 1 j_1 k_1) \quad (3 \quad 1 q_n S_n)^{-1/2} t^{1/2} \\
& fY_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n + \prod_{j_2 \in T} \prod_{k_2=1} \prod_{j_1 \in T} \prod_{k_1=1} \prod_{i_1 \in H_1} P \quad n_1^{-1} \times \prod_{j_2 k_2}^{-1/2} (i_1 j_2 k_2 \quad 1 j_2 k_2) \\
& (3 \quad 1 q_n S_n)^{-1/2} t^{1/2} fY_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \\
& 4(q_n S_n)^2 \exp(-c_4 n q_n^{-1} S_n^{-1} t);
\end{aligned}$$

for some universal constant $c_4 > 0$, where the last inequality follows from Hoeffding inequality and the definition of \mathcal{M}_n . Therefore, we set $c_5 = 3c_4^{-1}$ and plug $t = c_5 q_n S_n \log(q_n S_n) = n$ into the above inequality to obtain

$$P \Omega_3 \quad c_5 q_n S_n \log(q_n S_n) = n fY_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \quad 4(q_n S_n)^{-1};$$

Similar reasoning leads to

$$P \Omega_4 \quad c_6 q_n S_n \log(q_n S_n) = n fY_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \quad 4(q_n S_n)^{-1};$$

for some universal constant $c_6 > 0$. Accordingly, we set $c_7 = c_2 + c_3 + c_5 + c_6 + 1$. By combining the above two inequalities with (76), (77), and (75), it can be deduced that

$$\begin{aligned}
& P \quad k \Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} \quad \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} k_2 \quad c_7 f q_n^2 S_n^2 \log(q_n S_n) = n g^{1/2} fY_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \\
& 1 \quad 12(q_n S_n)^{-1}. \tag{78}
\end{aligned}$$

Finally, we have

$$\begin{aligned}
& P \left(k \Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} k_2 \leq c_7 \bar{f} q_n^2 S_n^2 \log(q_n S_n) = n g^{1/2} \right) \\
& \quad \times \\
& P \left(k \Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} k_2 \leq c_7 \bar{f} q_n^2 S_n^2 \log(q_n S_n) = n g^{1/2} \right) \\
& \quad \times \\
& P \left(f Y_i = y_i g_{i=1}^n \right) = P \left(f Y_i = y_i g_{i=1}^n \right) \\
& \quad \times \\
& \bar{f} \left(12(q_n S_n)^{-1} g \right) P \left(f Y_i = y_i g_{i=1}^n \right) = \bar{f} \left(12(q_n S_n)^{-1} g P(\mathcal{M}_n) \right) \\
& \quad \times \\
& \left(1 - 12 \bar{f}(q_n S_n)^{-1} + \exp(-n_1/12) + \exp(-n_2/12) \right) g;
\end{aligned}$$

where the second inequality is by (78), and the last inequality follows from Lemma 3. This completes the proof. \square

Lemma 8. Assume the following conditions (a)–(b):

(a) $q_n^2 S_n^2 \log(q_n S_n) = o(n)$.

(b) $c_1 \min(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \leq \max(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \leq c_2$, for some universal constants $0 < c_1 < c_2$.

Then we have the following properties:

1) There exist universal constants $c_3 > 0$ and $c_4 > 0$ such that

$$P(k \Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} k_2 \leq c_3) \geq 1 - c_4 \bar{f}(q_n S_n)^{-1} + \exp(-n_1/12) + \exp(-n_2/12)g.$$

2) There exist universal constants $c_5 > 0$ and $c_6 > 0$ such that

$$P(k \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} k_2 \leq c_5) \geq 1 - c_6 \bar{f}(q_n S_n)^{-1} + \exp(-n_1/12) + \exp(-n_2/12)g.$$

Proof of Lemma 8: First of all, we note that

$$k \Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} k_2 \leq k \Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2} k_2 + c_2;$$

where c_2 is defined in condition (b). Together with condition (a) and Lemma 7, it can be concluded that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that with probability at least $1 - c_3 \bar{f}(q_n S_n)^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})g$,

$$k\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} k_2 \leq c_4 \bar{f} q_n^2 S_n^2 \log(q_n S_n) = n g^{1/2} + c_2 \leq 2c_2;$$

which completes the proof of property 1). To show the second property, we first notice that

$$k\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} k_2 = \frac{-1}{\min(\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2})}. \quad (79)$$

Moreover, it is apparent to deduce that

$$\min(\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \leq c_1 \leq k\Lambda_T^{(1)-1/2} S_{TT}^{(1)} \Lambda_T^{(1)-1/2} \leq \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}$$

$$\begin{aligned}
2) \quad & P \quad f \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right)' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right) g = f \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right)' \hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right) g \\
1 \quad & c_3 \quad f \log(q_n s_n) = n g^{1/2} + f \log \log(n) = n g^{1/2} \\
1 \quad & c_4 \quad (q_n s_n)^{-1} + f \log(n) g^{-1} + \exp(-n_1 = 12) + \exp(-n_2 = 12) .
\end{aligned}$$

Proof of Lemma 9: First of all, we note that

$$\operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right)' \hat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right) = \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right)' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right) + \Omega_1 + 2\Omega_2; \quad (80)$$

where

$$\begin{aligned}
\Omega_1 &= \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right)' (\hat{\Lambda}_T^{(1)1/2} \Lambda_T^{(1)-1/2} \quad l_{q_n s_n}) (\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2}) (\Lambda_T^{(1)-1/2} \hat{\Lambda}_T^{(1)1/2} \quad l_{q_n s_n}) \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right); \\
\Omega_2 &= \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right)' (\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2}) (\Lambda_T^{(1)-1/2} \hat{\Lambda}_T^{(1)1/2} \quad l_{q_n s_n}) \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right);
\end{aligned}$$

For the term Ω_1 , it can be deduced that

$$\Omega_1 \quad q_n s_n k \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} k_2 \quad k \hat{\Lambda}_T^{(1)1/2} \Lambda_T^{(1)-1/2} \quad l_{q_n s_n} k_{\max}^2;$$

Together with condition (a), condition (b), Lemma 8, and Lemma 5, it can be concluded that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that

$$P f \Omega_1 \quad c_3 q_n s_n \log(q_n s_n) = n g \quad 1 \quad c_4 f(q_n s_n)^{-1} + \exp(-n_1 = 12) + \exp(-n_2 = 12) g; \quad (81)$$

For the term Ω_2 , one has

$$\begin{aligned}
j \Omega_2 j & \quad k (\Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2}) \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right) k_1 \quad k (\Lambda_T^{(1)-1/2} \hat{\Lambda}_T^{(1)1/2} \quad l_{q_n s_n}) \operatorname{sgn} \left(\begin{smallmatrix} (1) \\ T \end{smallmatrix} \right) k_{\infty} \\
& \quad q_n s_n k \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} k_2 \quad k \hat{\Lambda}_T^{(1)1/2} \Lambda_T^{(1)-1/2} \quad l_{q_n s_n} k_{\max};
\end{aligned}$$

Together with condition (a), condition (b), Lemma 8, and Lemma 5, it can be deduced that there exist universal constants $c_5 > 0$ and $c_6 > 0$ such that

$$P [j \Omega_2 j \quad c_5 f q_n^2 s_n^2 \log(q_n s_n) = n g^{1/2}] \quad 1 \quad c_6 f(q_n s_n)^{-1} + \exp(-n_1 = 12) + \exp(-n_2 = 12) g;$$

Together with (80) and (81), it can be concluded that there exist universal constants $c_7 > 0$ and $c_8 > 0$ such that with probability at least $1 - c_7 f(q_n S_n)^{-1} + \exp(-n^{-1}) + \exp(-n^{-2})g$,

$$\begin{aligned} & \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g = \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g - 1 \\ & c_8 f q_n^2 S_n^2 \log(q_n S_n) = n g^{1/2}. \end{aligned}$$

Moreover, we note that

$$\begin{aligned} & \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g = \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g - 1 \\ & \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g^{-1} \\ & \text{sgn}(\mathbf{T}^{(1)})' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) - \text{sgn}(\mathbf{T}^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) + \\ & \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \Lambda_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g = \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g - 1 \\ & c_2 (q_n S_n)^{-1} \text{sgn}(\mathbf{T}^{(1)})' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) - \text{sgn}(\mathbf{T}^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) \\ & + \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \Lambda_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g = \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g - 1 ; \end{aligned}$$

where the last inequality is based on condition (b). Therefore, by combining Lemma 22 with the above two inequalities, we conclude that there exist universal constants $c_9 > 0$ and $c_{10} > 0$ such that with probability at least $1 - c_9 (q_n S_n)^{-1} + f \log(n) g^{-1} + \exp(-n^{-1}) + \exp(-n^{-2})$,

$$\begin{aligned} & \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \hat{\Lambda}_T^{(1)1/2} \mathcal{S}_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g = \bar{f} \text{sgn}(\mathbf{T}^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\mathbf{T}^{(1)}) g - 1 \\ & c_{10} f \log(q_n S_n) = n g^{1/2} + q_n S_n = n + f \log \log(n) = n g^{1/2} \\ & 2c_{10} f \log(q_n S_n) = n g^{1/2} + f \log \log(n) = n g^{1/2} ; \end{aligned}$$

which completes the proof of property 1). To show the second property, we notice the fact

that

$$\begin{aligned}
& \widehat{f} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) g = \widehat{f} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix})' \widehat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \widehat{\Lambda}_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) g \quad 1 \\
& = \widehat{f} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix})' \widehat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \widehat{\Lambda}_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) g = \widehat{f} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) g \quad 1 \\
& \widehat{f} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix})' \widehat{\Lambda}_T^{(1)1/2} S_{TT}^{(1)-1} \widehat{\Lambda}_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) g = \widehat{f} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) g^{-1}.
\end{aligned}$$

Together with property 1), property 2) follows directly, which finishes the proof. \square

Lemma 10. *Assume the following conditions (a){(b):*

$$(a) \quad q_n s_n = o(n).$$

$$(b) \quad c_1 \leq \min(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \leq \max(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \leq c_2, \text{ for some universal constants } 0 < c_1 < c_2.$$

Then there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that with probability at least

$$1 - c_3[(q_n s_n)^{-1} + \ell \log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})], \text{ we have:}$$

$$\begin{aligned}
& j_T^{\wedge(1)'} S_{TT}^{(1)-1} \widehat{\Lambda}_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) - j_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) j \\
& c_4 [q_n s_n = n + \ell \log(q_n s_n) = n g^{1/2} + \ell \log \log(n) = n g^{1/2}] j_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) j \\
& + c_4 (q_n s_n)^{1/2} \ell \log \log(n) = n g^{1/2} [1 + j_T^{(1)'} \Sigma_{TT}^{(1)-1} j_T^{(1)} + \widehat{f} j_T^{(1)'} \Sigma_{TT}^{(1)-1} j_T^{(1)} \log \log(n) = n g^{1/2}]^{1/2} \\
& + c_4 (q_n s_n)^{1/2} \ell \log(q_n s_n) = n g^{1/2} \widehat{f} q_n s_n \log(q_n s_n \log n) = n g^{1/2} \\
& [1 + j_T^{(1)'} \Sigma_{TT}^{(1)-1} j_T^{(1)} + \widehat{f} j_T^{(1)'} \Sigma_{TT}^{(1)-1} j_T^{(1)} \log \log(n) = n g^{1/2}]^{1/2}.
\end{aligned}$$

Proof of Lemma 10: First of all, we note that

$$j_T^{\wedge(1)'} S_{TT}^{(1)-1} \widehat{\Lambda}_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) - j_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) j = \Omega_1 + \Omega_2; \quad (82)$$

where

$$\begin{aligned}
\Omega_1 &= j_T^{\wedge(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) - j_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) j; \\
\Omega_2 &= j_T^{\wedge(1)'} S_{TT}^{(1)-1} \widehat{\Lambda}_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) - j_T^{\wedge(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\begin{smallmatrix} (1) \\ T \end{smallmatrix}) j;
\end{aligned}$$

For the term Ω_1 , Lemma 21 together with condition (b) imply that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that with probability at least $1 - c_3[\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})]$,

$$\Omega_1 \leq c_4 f q_n S_n \log \log(n) = n g^{1/2}. \quad (83)$$

For the term Ω_2 , it is clear that

$$\Omega_2 = \Pi_1 + \Pi_2; \quad (84)$$

where

$$\Pi_1 = j_T^{(1)'} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} (\Lambda_T^{(1)-1/2} \hat{\Lambda}_T^{(1)1/2} - I_q)$$

$$1 - c_7[(q_n s_n)^{-1} + f \log(n) g^{-1} + \exp(-n^{-1}) + \exp(-n^{-2})],$$

$$\begin{aligned} \Pi_1 &= c_8 f \log(q_n s_n) = n g^{1/2} \int_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) j + \\ &= c_8 \hat{f} q_n s_n \log(q_n s_n) = n g^{1/2} \hat{f} q_n s_n \log(q_n s_n \log n) = n g^{1/2} \\ &= [1 + \int_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{f}_T^{(1)} + \int_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{f}_T^{(1)} \log \log(n) = n g^{1/2}]^{1/2}; \end{aligned} \quad (85)$$

To bound the term Π_2 , we note that

$$\Pi_2 = c_1^{-1} q_n s_n (1 + \Upsilon_1) \int_T \Upsilon_2 j + \int_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) j + \Upsilon_3 g^{-1} \Upsilon_1; \quad (86)$$

where c_1 is defined in condition (b) and

$$\begin{aligned} \Upsilon_1 &= \int_T \hat{f} \text{sgn}(\hat{f}_T^{(1)}) \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) g \hat{f} \text{sgn}(\hat{f}_T^{(1)}) \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) g^{-1} - 1 j; \\ \Upsilon_2 &= \int_T \hat{f} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) g \hat{f} \text{sgn}(\hat{f}_T^{(1)}) \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) g^{-1} \\ &= \int_T \hat{f} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) g \hat{f} \text{sgn}(\hat{f}_T^{(1)}) \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) g^{-1}; \\ \Upsilon_3 &= \int_T \hat{f} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) - \int_T \hat{f} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) j; \end{aligned}$$

For the term Υ_1 , Lemma 22 entails that there exist universal constants $c_9 > 0$ and $c_{10} > 0$ such that with probability at least $1 - c_9[f \log(n) g^{-1} + \exp(-n^{-1}) + \exp(-n^{-2})]$,

$$\Upsilon_1 \leq c_{10} [q_n s_n = n + f \log \log(n) = n g^{1/2}]; \quad (87)$$

For the term Υ_2 , by using similar arguments as in the proof of Lemma 23, it can be deduced that there exist universal constants $c_{11} > 0$ and $c_{12} > 0$ such that conditional on any nonempty $\hat{f} Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_T g$, and for any $t \geq 0$,

$$\begin{aligned} P \int_T \Upsilon_2 j \leq t \hat{f} Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_T g \\ c_{11} \exp(-c_{12} n \hat{f} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \hat{f}_T g^{-1} \hat{f} \text{sgn}(\hat{f}_T^{(1)}) \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}_T^{(1)}) g t^2); \end{aligned}$$

By plugging $t = c_{13} \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f} g^{1/2} \hat{f} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}}) g^{-1/2}) \hat{f} \log \log(n) = n g^{1/2}$ with $c_{13} = c_{12}^{-1/2}$ into the above inequality, it yields that

$$\begin{aligned}
& P \int \Upsilon_{2j} \quad c_{13} \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f} g^{1/2} \hat{f} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}}) g^{-1/2}) \\
& \quad \hat{f} \log \log(n) = n g^{1/2} \quad \hat{f} Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \\
& \quad 1 \quad c_{11} \hat{f} \log(n) g^{-1}: \tag{88}
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
& P \int \Upsilon_{2j} \quad c_{13} \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f} g^{1/2} \hat{f} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}}) g^{-1/2}) \\
& \quad \hat{f} \log \log(n) = n g^{1/2} \\
& \quad \times \\
& \quad P \int \Upsilon_{2j} \quad c_{13} \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f} g^{1/2} \hat{f} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}}) g^{-1/2}) \\
& \quad \{y_i\}_{i=1}^n \in \mathcal{M}_n \\
& \quad \hat{f} \log \log(n) = n g^{1/2} \quad \hat{f} Y_i = y_i g_{i=1}^n \quad P \hat{f} Y_i = y_i g_{i=1}^n \\
& \quad \times \quad \cap^Z \\
& = \\
& \quad P \int \Upsilon_{2j} \quad c_{13} \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f} g^{1/2} \hat{f} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}}) g^{-1/2}) \\
& \quad \{y_i\}_{i=1}^n \in \mathcal{M}_n \quad \hat{v}_T \\
& \quad \hat{f} \log \log(n) = n g^{1/2} \quad \hat{f} Y_i = y_i g_{i=1}^n \setminus \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \quad \hat{f}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}}) g^{-1/2}) \\
& \quad \times \\
& \quad [1 \quad c_{11} \hat{f} \log(n) g^{-1}] \quad P \hat{f} Y_i = y_i g_{i=1}^n = [1 \quad c_{11} \hat{f} \log(n) g^{-1}] \quad P(\mathcal{M}_n) \\
& \quad \{y_i\}_{i=1}^n \in \mathcal{M}_n \\
& \quad 1 \quad c_{14} \hat{f} \log(n) g^{-1} + \exp(-n_{1=12}) + \exp(-n_{2=12}) ;
\end{aligned}$$

for some universal constant $c_{14} > 0$, where $\hat{f}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}}) g^{-1/2}) \hat{f} Y_i = y_i g_{i=1}^n$ denotes the conditional density function, and the second inequality is by (88). Together with Lemma 19 yields the result that there exist universal constants $c_{15} > 0$ and $c_{16} > 0$ such that with probability at least $1 - c_{15}[\hat{f} \log(n) g^{-1} + \exp(-n_{1=12}) + \exp(-n_{2=12})]$,

$$\begin{aligned}
& \int \Upsilon_{2j} \quad c_{16} [q_n S_n = n + \log \log(n) = n + \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{f}'_{T \Sigma_{TT}^{(1)-1}}) \log \log(n) = n g^{1/2}]^{1/2} \\
& \quad (q_n S_n)^{-1/2} \hat{f} \log \log(n) = n g^{1/2}: \tag{89}
\end{aligned}$$

For the term Υ_3 , Lemma 21 leads to the result that there exist universal constants $c_{17} > 0$ and $c_{18} > 0$ such that with probability at least $1 - c_{17}[\mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})]$,

$$\Upsilon_3 \leq c_{18} \mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})).$$

Together with (87), (89) and (86), it can be observed that there exist universal constants $c_{19} > 0$ and $c_{20} > 0$ such that with probability at least $1 - c_{19}[\mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})]$,

$$\begin{aligned} \Pi_2 &\leq c_{20}[\mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2}))^{1/2} \\ &\quad (\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2}))^{1/2} + c_{20}[\mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})) \\ &\quad + c_{20} \mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2}))^{1/2}]. \end{aligned}$$

Together with (84) and (85), there exist universal constants $c_{21} > 0$ and $c_{22} > 0$ such that with probability at least $1 - c_{21}[\mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})]$,

$$\begin{aligned} \Omega_2 &\leq c_{22}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2}))^{1/2} \\ &\quad [\mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2}))^{1/2} \\ &\quad + c_{22}[\mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})) \\ &\quad + c_{22} \mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2}))^{1/2} \\ &\quad + c_{22} \mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2}))^{1/2} \\ &\quad [1 + \mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2}))^{1/2}]. \end{aligned}$$

Together with (82) and (83), it can be concluded that there exist universal constants $c_{23} > 0$ and $c_{24} > 0$ such that with probability at least $1 - c_{23}[\mathbb{P}(\log(n)g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2})]$,

$$\exp(-n_1=12) + \exp(-n_2=12)],$$

$$\begin{aligned} & j_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn}(\Lambda_T^{(1)}) j_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\Lambda_T^{(1)}) j \\ & c_{24} [q_n s_n = n + f \log(q_n s_n) = ng^{1/2} + f \log \log(n) = ng^{1/2}] j_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\Lambda_T^{(1)}) j \\ & + c_{24} (q_n s_n)^{1/2} f \log \log(n) = ng^{1/2} [1 + \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)-1/2} + f \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)-1/2} \log \log(n) = ng^{1/2}]^{1/2} \\ & + c_{24} (q_n s_n)^{1/2} f \log(q_n s_n) = ng^{1/2} f q_n s_n \log(q_n s_n \log n) = ng^{1/2} \\ & [1 + \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)-1/2} + f \Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)-1/2} \log \log(n) = ng^{1/2}]^{1/2}; \end{aligned}$$

which completes the proof. \square

Lemma 11. *Assume the following conditions (a)-(d):*

$$(a) \max f q_n^2 s_n^2 \log(q_n s_n); q_n s_n \log(p_n - q_n) g = o(n).$$

$$(b) c_1 \min(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \max(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2}) \leq c_2, \text{ for some universal constants } 0 < c_1 < c_2.$$

$$(c) K_1 \log f(p_n - q_n) s_n \log ng = (n \binom{2}{n}) \prod_{j \in T} \prod_{k=1}^{s_n} \frac{1}{j k} \leq K_2 \log f(p_n - q_n) s_n \log ng = (n \binom{2}{n}) \leq 1, \text{ for some sufficiently large universal constants } K_2 > K_1 > 0.$$

$$\begin{aligned} (d) \min_{j \in T} \min_{k \leq s_n} \frac{1}{j k} > K_3 [\log f(p_n - q_n) s_n \log ng = (n \binom{2}{n})]^{1/2} f \log(q_n s_n \log n) = ng^{1/2} + \\ K_3 [\log f(p_n - q_n) s_n \log ng = (n \binom{2}{n})] k \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\Lambda_T^{(1)}) k_\infty + K_3 [\log f(p_n - q_n) s_n \log ng = (n \binom{2}{n})] \\ [f q_n s_n \log(q_n s_n) = ng^{1/2} + f q_n s_n \log \log(n) = ng^{1/2}], \text{ for some sufficiently large universal constant } K_3 > 0. \end{aligned}$$

Then there exists a universal constant $c_3 > 0$ such that:

$$P \sum_{k \in S_1} \left[\log f(\rho_n) - \log(q_n S_n) \right] \leq c_3 \left[\log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/3}) \right];$$

where $\hat{\rho}_T^{(1)} = S_{TT}^{(1)-1} \hat{\rho}_T^{(1)}$, and also recall that \tilde{v}_T is defined in Lemma 2.

Proof of Lemma 11: First of all, we denote the two index sets S_1 and S_2 as

$$S_1 = \{k : \hat{\rho}_k^{(1)} \sum_{TT}^{(1)-1} \hat{\rho}_T^{(1)} > 0g\}; \quad S_2 = \{k : \hat{\rho}_k^{(1)} \sum_{TT}^{(1)-1} \hat{\rho}_T^{(1)} < 0g\};$$

By definition, we have $S_1 \cap S_2 = \emptyset$. Moreover, by using Lemma 23 and conditions (a)–(c), it can be shown that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that

$$\begin{aligned} P \sum_{k \in S_1} \left[\log f(\rho_n) - \log(q_n S_n) \right] &\leq c_3 \left[\log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/3}) \right] \\ &\leq c_3 \left[\log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/3}) \right] \\ &\leq c_4 \left[\log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/3}) \right]; \end{aligned} \quad (90)$$

For the term $\hat{\rho}_T^{(1)} \sum_{TT}^{(1)-1} \hat{\rho}_T^{(1)}$, conditions (b) and (c) entail that

$$\hat{\rho}_T^{(1)} \sum_{TT}^{(1)-1} \hat{\rho}_T^{(1)} \leq \log f(\rho_n) - \log(q_n S_n) \leq (n^{-2})^{-1/2}; \quad \text{as } n \rightarrow \infty;$$

Together with (90), there exist positive universal constants c_5, c_6 and c_7 such that

$$\begin{aligned} P \sum_{k \in S_1} \left[\log f(\rho_n) - \log(q_n S_n) \right] &\leq c_5 \left[\log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/3}) \right] \\ &\leq c_6 \left[\log f(\rho_n) - \log(q_n S_n) \right]^{1/2} \log(n) g^{-1} \\ &\leq c_7 \left[\log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/3}) \right]; \end{aligned} \quad (91)$$

By choosing $K_3 > c_6 = c_5$ in condition (d), (91) together with condition (d) further implies that

$$P \bigcap_{k \in \mathcal{S}_1} e_k' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)} > 0 \quad 1 - c_7[\ell \log(n)g^{-1} + \exp(-n_1/2) + \exp(-n_2/2)]:$$

Likewise, it can be deduced that there exists a universal constant $c_8 > 0$ such that

$$P \bigcap_{k \in \mathcal{S}_2} e_k' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)} < 0 \quad 1 - c_8[\ell \log(n)g^{-1} + \exp(-n_1/2) + \exp(-n_2/2)]:$$

Putting the above two inequalities together implies that there exists a universal constant $c_9 > 0$ such that

$$P \tilde{f} \operatorname{sgn}(\hat{\Lambda}_T^{(1)}) = \operatorname{sgn}(\hat{\Lambda}_T^{(1)})g \quad 1 - c_9[\ell \log(n)g^{-1} + \exp(-n_1/2) + \exp(-n_2/2)]: \quad (92)$$

Moreover, it can be recalled from Lemma 2 that the quantity \tilde{v}_T can be formulated as

$$\tilde{v}_T = \hat{\#} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)} \operatorname{sgn}(\hat{\Lambda}_T^{(1)});$$

where

$$\begin{aligned} \hat{\#} &= \ell n_1 n_2 n^{-1} (n_2)^{-1} g \mathbb{1} + \hat{\Lambda}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \operatorname{sgn}(\hat{\Lambda}_T^{(1)})g \\ &\quad 1 + \ell n_1 n_2 n^{-1} (n_2)^{-1} g \hat{\Lambda}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)-1}. \end{aligned}$$

To this end, by combining conditions (a)–(c) with Lemma 10, it can be deduced that there exists a universal constant $c_{10} > 0$ such that with probability at least $1 - c_{10}[(q_n S_n)^{-1} + \ell \log(n)g^{-1} + \exp(-n_1/2) + \exp(-n_2/2)]$,

$$\hat{\Lambda}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \operatorname{sgn}(\hat{\Lambda}_T^{(1)}) = \hat{\Lambda}_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\hat{\Lambda}_T^{(1)}) \mathbb{1} + o(1)g + o(1):$$

Similarly, by combining conditions (a)–(c) with Lemma 4, it can be deduced that there exists a universal constant $c_{11} > 0$ such that with probability at least $1 - c_{11}[\ell \log(n)g^{-1} + \exp(-n_1/2) + \exp(-n_2/2)]$,

$$\hat{\Lambda}_T^{(1)'} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)} = \hat{\Lambda}_T^{(1)'} \Sigma_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)} \mathbb{1} + o(1)g:$$

According to the above three inequalities and Lemma 3, it can be concluded that there exist universal constants $c_{12} > 0$ and $c_{13} > 0$ such that with probability at least $1 - c_{12}[(q_n S_n)^{-1} + f \log(n)g^{-1} + \exp(-n_1/12) + \exp(-n_2/12)]$,

$$\hat{\#} \leq c_{13} \frac{1}{n} f + \frac{1}{n} \sum_{T \in \mathcal{T}} \Lambda_T^{(1)1/2} \text{sgn}(\frac{1}{T}) g f + \frac{1}{n} \sum_{T \in \mathcal{T}} \Lambda_T^{(1)1/2} g^{-1}.$$

For the term $\frac{1}{n} \sum_{T \in \mathcal{T}} \Lambda_T^{(1)1/2} \text{sgn}(\frac{1}{T})$, one has

$$\begin{aligned} & \frac{1}{n} \sum_{T \in \mathcal{T}} \Lambda_T^{(1)1/2} \text{sgn}(\frac{1}{T}) = \frac{1}{n} f \sum_{T \in \mathcal{T}} \Lambda_T^{(1)1/2} \text{sgn}(\frac{1}{T}) g^{1/2} \\ & f \text{sgn}(\frac{1}{T}) \Lambda_T^{(1)1/2} \sum_{T \in \mathcal{T}} \Lambda_T^{(1)1/2} \text{sgn}(\frac{1}{T}) g^{1/2} \lesssim \frac{1}{n} f q_n S_n \sum_{T \in \mathcal{T}} \Lambda_T^{(1)1/2} \text{sgn}(\frac{1}{T}) g^{1/2} \\ & \lesssim [q_n S_n \log f(\rho_n - q_n) S_n \log n g = n]^{1/2} \lesssim o(1); \end{aligned} \quad (93)$$

where the second and the third inequalities are based on (b) and (c), and the last inequality follows from (a). Piecing the above two inequalities together yields that there exist universal constants $c_{14} > 0$ and $c_{15} > 0$ such that with probability at least $1 - c_{14}[(q_n S_n)^{-1} + f \log(n)g^{-1} + \exp(-n_1/12) + \exp(-n_2/12)]$,

$$\hat{\#} \leq c_{15} \frac{1}{n} f \sum_{T \in \mathcal{T}} \Lambda_T^{(1)1/2} \text{sgn}(\frac{1}{T}) g^{-1}.$$

Together with (91) and (92), it can be deduced that there exist universal constants $c_{16}, c_{17}, c_{18} > 0$ such that

$$\begin{aligned} \mathcal{P} \setminus \bigcap_{k \in \mathcal{S}_1} & \hat{\#} \leq \frac{1}{n} \sum_{k \in \mathcal{S}_1} \Lambda_k^{(1)1/2} S_{TT}^{(1)1/2} \text{sgn}(\frac{1}{T}) \leq c_{17} \frac{1}{n} f \sum_{T \in \mathcal{T}} \Lambda_T^{(1)1/2} \text{sgn}(\frac{1}{T}) g^{-1} + c_{18} [\log f(\rho_n - q_n) S_n \log n g = n]^{1/2} f \log(q_n S_n \log n) = n g^{1/2} \\ & \leq c_{16} [(q_n S_n)^{-1} + f \log(n)g^{-1} + \exp(-n_1/12) + \exp(-n_2/12)]; \end{aligned}$$

In addition, utilizing Lemma 24 and conditions (a)–(c), it can also be justified that there

exist universal constants $c_{19} > 0$ and $c_{20} > 0$ such that

$$\begin{aligned}
& P \quad \text{h} \setminus \text{n} \\
& \quad \underset{k \in \mathcal{S}_1}{\sum} j \ell'_k \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn} \left(\binom{(1)}{T} \right) j \quad \underset{k \in \mathcal{S}_1}{\sum} j \ell'_k \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\binom{(1)}{T} \right) j \\
& \quad + c_{19} \underset{k \in \mathcal{S}_1}{\sum} [q_n s_n = n + f \log(q_n s_n \log n) = n g^{1/2}] j \ell'_k \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\binom{(1)}{T} \right) j \\
& \quad + c_{19} \underset{k \in \mathcal{S}_1}{\sum} [f q_n s_n \log(q_n s_n) = n g^{1/2} + f q_n s_n \log \log(n) = n g^{1/2}] \text{oi} \\
& \quad 1 \quad c_{20} [(q_n s_n)^{-1} + f \log(n) g^{-1} + \exp(-n_1 = 12) + \exp(-n_2 = 12)]:
\end{aligned}$$

Based on the above two inequalities, it is seen that there exist positive universal constants

c_{21} , c_{22} and c_{23} that

$$\begin{aligned}
& P \quad \text{h} \setminus \text{n} \\
& \quad \underset{k \in \mathcal{S}_1}{\sum} \ell'_k \Lambda_T^{(1)1/2} \tilde{V}_T \quad c_{21} f \binom{(1)}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T} g^{-1} \\
& \quad \ell'_k \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \binom{(1)}{T} \quad c_{22} [\log f(\rho_n \quad q_n) s_n \log n g = (n \quad n)^2]^{1/2} f \log(q_n s_n \log n) = n g^{1/2} \\
& \quad c_{22} [\log f(\rho_n \quad q_n) s_n \log n g = (n \quad n)] j \ell'_k \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\binom{(1)}{T} \right) j \\
& \quad c_{22} [\log f(\rho_n \quad q_n) s_n \log n g = (n \quad n)] [f q_n s_n \log(q_n s_n) = n g^{1/2} + f q_n s_n \log \log(n) = n g^{1/2}] \text{oi} \\
& \quad 1 \quad c_{23} [(q_n s_n)^{-1} + f \log(n) g^{-1} + \exp(-n_1 = 12) + \exp(-n_2 = 12)]:
\end{aligned}$$

By choosing $K_3 > c_{22}$ in condition (d), it follows from condition (d) and the above inequality

that

$$P \quad \text{h} \setminus \text{n} \quad \underset{k \in \mathcal{S}_1}{\sum} \ell'_k \Lambda_T^{(1)1/2} \tilde{V}_T > 0 \quad \text{oi} \quad 1 \quad c_{23} [(q_n s_n)^{-1} + f \log(n) g^{-1} + \exp(-n_1 = 12) + \exp(-n_2 = 12)]:$$

Similar reasoning leads to the result that there exists a universal constants $c_{24} > 0$ such

that

$$P \quad \text{h} \setminus \text{n} \quad \underset{k \in \mathcal{S}_2}{\sum} \ell'_k \Lambda_T^{(1)1/2} \tilde{V}_T < 0 \quad \text{oi} \quad 1 \quad c_{24} [(q_n s_n)^{-1} + f \log(n) g^{-1} + \exp(-n_1 = 12) + \exp(-n_2 = 12)]:$$

Based on (92) and the above two inequalities, there exists a universal constant $c_{25} > 0$ such

that

$$P \tilde{\text{sgn}}(\tilde{V}_T) = \text{sgn}\left(\frac{\cdot}{T}\right) = \text{sgn}\left(\frac{\hat{\cdot}}{T}\right)g$$

$$1 - c_{25}[(q_n S_n)^{-1} + \log(n)g^{-1} + \exp(-n^{-1}) + \exp(-n^{-2})];$$

which concludes the proof. \square

Lemma 12. Let a_n and b_n be any two sequences of constants such that $a_n \neq 1$ and $b_n \neq 0$.

Also let X_n and U_n be any two sequences of random variables such that $X_n = o_p(1)$ and $U_n = o_p(1)$. Assume that we have the following conditions (a)-(b):

$$(a) \ a_n X_n = o_p(1).$$

$$(b) \ a_n^{1/2}(U_n - b_n) = o_p(1).$$

Then we have the following property:

$$\Phi\left(a_n^{1/2}(1 + X_n) + U_n\right) = \Phi\left(a_n^{1/2} + b_n\right) + o_p(1);$$

Proof of Lemma 12: The proof is analogous to that of Lemma 1 in [Shao et al. \(2011\)](#). \square

Lemma 13. Consider a pair A, B of $p \times p$ matrices, assume the following condition (a):

$$(a) \ \min(A - B) \geq 0.$$

Then we have the following property:

$$\min(A) \leq \min(B); \quad \max(A) \leq \max(B);$$

Proof of Lemma 13: First of all, we have

$$\min(A) \leq \min(A - B) + \min(B) \leq \min(B);$$

where the last inequality is by condition (a). Similarly, we also have

$$\max(A) \leq \min(A \ B) + \max(B) \leq \max(B);$$

where the last inequality is by condition (a) as well, which completes the proof. \square

Lemma 14. For any $p \times p$ square matrix A , partitioned as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix};$$

where A_{11} is a $k \times k$ matrix for some positive integer $k < p$, assume we have the following condition (a):

$$(a) \ c_1 \leq \min(A) \leq \max(A) \leq c_2, \text{ for some universal constants } 0 < c_1 < c_2.$$

Then we have the following properties:

$$1) \ c_1 \leq \min(A_{11} \ A_{12}A_{22}^{-1}A_{21}) \leq \max(A_{11} \ A_{12}A_{22}^{-1}A_{21}) \leq c_2,$$

$$c_1 \leq \min(A_{22} \ A_{21}A_{11}^{-1}A_{12}) \leq \max(A_{22} \ A_{21}A_{11}^{-1}A_{12}) \leq c_2.$$

$$2) \ \max(A_{12}A_{22}^{-1}A_{21}) \leq \max(A_{11}) \leq c_2,$$

$$\max(A_{21}A_{11}^{-1}A_{12}) \leq \max(A_{22}) \leq c_2,$$

$$\min(A_{12}A_{22}^{-1}A_{21}) \leq \min(A_{11}),$$

$$\min(A_{21}A_{11}^{-1}A_{12}) \leq \min(A_{22}).$$

Proof of Lemma 14: Based on condition (a), we have

$$c_2^{-1} \leq \min(A^{-1}) \leq \max(A^{-1}) \leq c_1^{-1};$$

where A^{-1} can be expressed as

$$A^{-1} = \begin{pmatrix} (A_{11} \ A_{12}A_{22}^{-1}A_{21})^{-1} & A_{11}^{-1}A_{12}(A_{22} \ A_{21}A_{11}^{-1}A_{12})^{-1} \\ A_{22}^{-1}A_{21}(A_{11} \ A_{12}A_{22}^{-1}A_{21})^{-1} & (A_{22} \ A_{21}A_{11}^{-1}A_{12})^{-1} \end{pmatrix};$$

Hence, we have

$$\begin{aligned} c_2^{-1} &= \min((A_{11} \quad A_{12}A_{22}^{-1}A_{21})^{-1}) \quad \max((A_{11} \quad A_{12}A_{22}^{-1}A_{21})^{-1}) \quad c_1^{-1}; \\ c_2^{-1} &= \min((A_{22} \quad A_{21}A_{11}^{-1}A_{12})^{-1}) \quad \max((A_{22} \quad A_{21}A_{11}^{-1}A_{12})^{-1}) \quad c_1^{-1}; \end{aligned}$$

which implies that

$$\begin{aligned} c_1 &= \min(A_{11} \quad A_{12}A_{22}^{-1}A_{21}) \quad \max(A_{11} \quad A_{12}A_{22}^{-1}A_{21}) \quad c_2; \\ c_1 &= \min(A_{22} \quad A_{21}A_{11}^{-1}A_{12}) \quad \max(A_{22} \quad A_{21}A_{11}^{-1}A_{12}) \quad c_2; \end{aligned}$$

finishing the proof of property 1). Finally, by combining property 1) with Lemma 13, the assertion in property 2) follows immediately, which completes the proof. \square

Lemma 15. Let $fX_1; \dots; X_{n+m}g$ be a sample of random vectors in \mathbb{R}^p . Denote

$$\begin{aligned} S_1 &= \sum_{i=1}^n (X_i \quad \bar{X}_1)(X_i \quad \bar{X}_1)' = (n-1); \quad \bar{X}_1 = \frac{1}{n} \sum_{i=1}^n X_i; \\ S_2 &= \sum_{i=n+1}^{n+m} (X_i \quad \bar{X}_2)(X_i \quad \bar{X}_2)' = (m-1); \quad \bar{X}_2 = \frac{1}{m} \sum_{i=n+1}^{n+m} X_i; \\ S &= \sum_{i=1}^{n+m} (X_i \quad \bar{X})(X_i \quad \bar{X})' = (n+m-2); \quad \bar{X} = \frac{1}{n+m} \sum_{i=1}^{n+m} X_i; \\ S_{pool} &= f(n-1)S_1 + (m-1)S_2g = (n+m-2): \end{aligned}$$

Then we have the following property:

$$S = S_{pool} + nm(n+m)^{-1}(n+m-2)^{-1}(\bar{X}_1 \quad \bar{X}_2)(\bar{X}_1 \quad \bar{X}_2)';$$

Proof of Lemma 15: The term S can be decomposed as $S = I_1 + I_2$ with

$$\begin{aligned} I_1 &= \sum_{i=1}^n (X_i \quad \bar{X})(X_i \quad \bar{X})' = (n+m-2); \\ I_2 &= \sum_{i=n+1}^{n+m} (X_i \quad \bar{X})(X_i \quad \bar{X})' = (n+m-2); \end{aligned}$$

For the term l_1 , one has

$$l_1 = (n-1)(n+m-2)^{-1}S_1 + nm^2(n+m)^{-2}(n+m-2)^{-1}(\bar{X}_1 \quad \bar{X}_2)(\bar{X}_1 \quad \bar{X}_2)'$$

By symmetry, we also have

$$l_2 = (m-1)(n+m-2)^{-1}S_2 + mn^2(n+m)^{-2}(n+m-2)^{-1}(\bar{X}_1 \quad \bar{X}_2)(\bar{X}_1 \quad \bar{X}_2)'$$

Based on the above results, we conclude that $S = S_{pool} + nm(n+m)^{-1}(n+m-2)^{-1}(\bar{X}_1 \quad \bar{X}_2)(\bar{X}_1 \quad \bar{X}_2)'$, which finishes the proof. \square

Lemma 16. Recall that $T = f\mathbf{1}; \dots; q_n g$. Assume the matrix $\Sigma_{TT}^{(1)}$ is invertible and consider the following optimization problem:

$$\min_{w_T \in \mathbb{R}^{qn}} \frac{1}{2} w_T' \Sigma_{TT}^{(1)} w_T + \frac{1}{2} w_T' \Lambda_T^{(1)1/2} w_T + n(\Lambda_T^{(1)1/2} w_T)' \text{sgn}(w_T);$$

where $w_T = (w_1'; \dots; w_{q_n}')'$ with sub-vectors $w_j = (w_{j1}; \dots; w_{js_n})' \in \mathbb{R}^{s_n}$. Let \tilde{w}_T be the solution of the optimization problem where $\tilde{w}_T = (\tilde{w}_1'; \dots; \tilde{w}_{q_n}')'$ with sub-vectors $\tilde{w}_j = (\tilde{w}_{j1}; \dots; \tilde{w}_{js_n})' \in \mathbb{R}^{s_n}$, then we have:

$$\tilde{w}_T = \frac{1}{2} (1 + n\Lambda_T^{(1)1/2} k_1) (1 + \frac{1}{2} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2})^{-1} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(w_T);$$

Proof of Lemma 16: First of all, based on first order condition, one has

$$\Sigma_{TT}^{(1)} w_T + \Lambda_T^{(1)1/2} w_T = \Lambda_T^{(1)1/2} \text{sgn}(w_T); \quad (94)$$

Moreover, according to Sherman-Morrison-Woodbury formula, we have

$$\begin{aligned} \Sigma_{TT}^{(1)} + \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} &= \Sigma_{TT}^{(1)-1} \left(\Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} + \Lambda_T^{(1)1/2} \right) \Sigma_{TT}^{(1)-1} \\ &= \Sigma_{TT}^{(1)-1} \left(1 + \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \right) \Sigma_{TT}^{(1)-1}. \end{aligned}$$

Finally, by combining the above two equations, we have

$$\begin{aligned}
\tilde{W}_T &= \Sigma_{TT}^{(1)} + \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)'} \Sigma_{TT}^{(1)-1} f \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)} \text{sgn}(\Lambda_T^{(1)}) g \\
&= f \Sigma_{TT}^{(1)-1} \frac{1}{2} (1 + \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)})^{-1} \Sigma_{TT}^{(1)-1} \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)-1} g \\
&\quad f \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)} \text{sgn}(\Lambda_T^{(1)}) g \\
&= f \Sigma_{TT}^{(1)-1} \frac{1}{2} (1 + \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)} \Lambda_T^{(1)})^{-1} \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)-1} g \\
&\quad f \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)} \text{sgn}(\Lambda_T^{(1)}) g \\
&= f \frac{1}{2} \Sigma_{TT}^{(1)-1} \frac{1}{2} (1 + \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)} \Lambda_T^{(1)})^{-1} \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)-1} g \\
&\quad f \frac{1}{2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)} \Sigma_{TT}^{(1)} \text{sgn}(\Lambda_T^{(1)}) \frac{1}{2} (1 + \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)} \Lambda_T^{(1)})^{-1} \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)-1} g \\
&= \frac{1}{2} (1 + \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)})^{-1} \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)} \Sigma_{TT}^{(1)-1} \text{sgn}(\Lambda_T^{(1)});
\end{aligned}$$

which finishes the proof. \square

Lemma 17. Consider the following optimization problem:

$$\min_{w \in \mathbb{R}^{p_n s_n}} \frac{1}{2} w' \Sigma^{(1)} + \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)'} \Sigma_{TT}^{(1)-1} w + \sum_{j=1}^n k \Lambda_j^{(1)/2} w_j k_1; \quad (95)$$

where $w = (w_1' \dots w_{p_n}')'$ with vectors $w_j = (w_{j1} \dots w_{js_n})' \in \mathbb{R}^{s_n}$. Assume we have the following conditions (a)-(c):

(a) $\Sigma_{TT}^{(1)}$ is invertible.

(b) $\frac{1}{2} (1 + \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)})^{-1} (\min_{j \in T} \min_{k \leq s_n} \frac{1}{2} j_{jk} k_1) > \frac{1}{2} \Lambda_T^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)} \Sigma_{TT}^{(1)} \text{sgn}(\Lambda_T^{(1)}) k_\infty$.

(c) $k \Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)/2} \text{sgn}(\Lambda_T^{(1)}) k_\infty \leq 1$, for a universal constant $\leq 2(0; 1]$.

Denote \hat{w} as $\hat{w} = (\hat{w}_T'; \hat{w}_N')' = (\tilde{w}_T'; 0)'$ with $\hat{w}_N = 0 \in \mathbb{R}^{(p_n - q_n) s_n}$, and $\hat{w}_T = \tilde{w}_T$ where \tilde{w}_T is defined in Lemma 16. Then we have the following properties:

1) \hat{w} is a global minimum of (95).

$$2) \operatorname{sgn}(\hat{w}) = \operatorname{sgn}(\hat{w}^{(1)}).$$

Proof of Lemma 17: First of all, based on (a), (b) and the definition of \hat{w} , it is trivial to deduce that $\operatorname{sgn}(\hat{w}) = \operatorname{sgn}(\hat{w}^{(1)})$, finishing the proof of 2). Moreover, according to the optimization theory, we know that \hat{w} is a global minimum of (95) if and only if

$$\Sigma_{TT}^{(1)} + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) & (1)' \\ T & T \end{bmatrix} \tilde{W}_T = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) \\ T \end{bmatrix} \Lambda_T^{(1)1/2} \operatorname{sgn}(\hat{w}^{(1)}); \quad (96)$$

$$k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) & (1)' \\ N & T \end{bmatrix} \tilde{W}_T \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) \\ N \end{bmatrix} k_\infty \quad n; \quad (97)$$

where (96) and (97) serve as the Karush-Kuhn-Tucker conditions. It is apparent that (96) follows from (94). In addition, observe that

$$\begin{aligned} & k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) & (1)' \\ N & T \end{bmatrix} \tilde{W}_T \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) \\ N \end{bmatrix} k_\infty \\ &= k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \begin{bmatrix} (1) & (1)' \\ T & T \end{bmatrix} \tilde{W}_T \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \begin{bmatrix} (1) \\ T \end{bmatrix} k_\infty \\ &= k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \quad l + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) & (1)' \\ T & T \end{bmatrix} \tilde{W}_T \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) \\ T \end{bmatrix} k_\infty; \end{aligned}$$

where the first and the second equalities follow from (10) in the main paper. For the term

$l + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) & (1)' \\ T & T \end{bmatrix} \tilde{W}_T$, we have

$$\begin{aligned} & l + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) & (1)' \\ T & T \end{bmatrix} \tilde{W}_T \\ &= \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} (1 + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \Lambda_T^{(1)1/2} \begin{bmatrix} (1) & (1)' \\ T & T \end{bmatrix} k_1) \begin{bmatrix} (1) \\ T \end{bmatrix} \Lambda_T^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\hat{w}^{(1)}) \\ & \quad \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) & (1)' \\ T & T \end{bmatrix} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\hat{w}^{(1)}) \\ &= \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) \\ T \end{bmatrix} \Lambda_T^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\hat{w}^{(1)}); \end{aligned}$$

where the first equality is by Lemma 16. To this end, based on the above two equations, we deduce that

$$\begin{aligned} & k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) & (1)' \\ N & T \end{bmatrix} \tilde{W}_T \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} (1) \\ N \end{bmatrix} k_\infty \\ &= \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} k\Lambda_N^{(1)-1/2} \Sigma_{NT}^{(1)} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \operatorname{sgn}(\hat{w}^{(1)}) k_\infty \quad n; \end{aligned}$$

where the last inequality is based on condition (c). According to the above results, it can be concluded that \hat{W} is a global minimum of (95), which completes the proof. \square

Lemma 18. For any $\% \geq (e^{-n/100}; 1=100)$, define the event $\mathcal{M}_{1n}(\%)$ as

$$\mathcal{M}_{1n}(\%) = \bigcap_{\substack{2^{-1}(q_n s_n = n) \\ 2(q_n s_n = n) + 16 \log(\%^{-1}) = ng^{1/2}}} \left(\begin{matrix} \text{Tr}(\hat{S}_{TT}^{(1)-1} \hat{\Sigma}_{TT}^{(1)}) \\ \text{Tr}(\hat{\Sigma}_{TT}^{(1)-1} \hat{S}_{TT}^{(1)}) \end{matrix} \right) \leq 1 + ng^{1/2} : \quad (97)$$

Assume the condition (a):

$$(a) \quad q_n s_n = o(n).$$

Then we have the following property:

$$P(\mathcal{M}_{1n}(\%)) \geq 1 - 2\% \geq 2 \exp(-n) - 2 \exp(-n/2); \quad \% \geq (e^{-n/100}; 1=100).$$

Proof of Lemma 18: First of all, based on condition (a) and the definition, it is clear to observe that conditional on any nonempty set $\mathcal{G}_i = \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n$, we have

$$(n-2) \text{Tr}(\hat{S}_{TT}^{(1)}) \mathcal{G}_i = \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \sim \text{Wishart}(n-2, \hat{\Sigma}_{TT}^{(1)}); \quad (98)$$

where the degree of freedom of the Wishart distribution is equal to $n-2$. Moreover, it is trivial to verify that conditional on $\mathcal{G}_i = \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n$, one has the fact that $\text{Tr}(\hat{S}_{TT}^{(1)}) \mathcal{G}_i = \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n$ is independent of $(n-2) \text{Tr}(\hat{S}_{TT}^{(1)}) \mathcal{G}_i = \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n$. Together with (98), condition (a) and Theorem 3.2.12 in Muirhead (1982), we reach a conclusion that

$$(n-2) \left(\begin{matrix} \text{Tr}(\hat{\Sigma}_{TT}^{(1)-1} \hat{S}_{TT}^{(1)}) \\ \text{Tr}(\hat{S}_{TT}^{(1)-1} \hat{\Sigma}_{TT}^{(1)}) \end{matrix} \right)^{-1} \mathcal{G}_i = \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \sim \chi^2_{n-q_n s_n - 1}.$$

Together with (A.2) and (A.3) in Johnstone and Lu (2009), we conclude that for any $t \geq [0; 1=2)$,

$$P \left(\begin{matrix} (n-q_n s_n - 1)^{-1} (n-2) \left(\begin{matrix} \text{Tr}(\hat{\Sigma}_{TT}^{(1)-1} \hat{S}_{TT}^{(1)}) \\ \text{Tr}(\hat{S}_{TT}^{(1)-1} \hat{\Sigma}_{TT}^{(1)}) \end{matrix} \right)^{-1} \\ \text{Tr}(\hat{S}_{TT}^{(1)}) \mathcal{G}_i = \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \end{matrix} \right) \leq 2 \exp(-t^2/3) (n-q_n s_n - 1)^{-1} t^2 = 16g:$$

For any $\% \geq (e^{-n/100}; 1=100)$, we plug $t = \sqrt{16(n - q_n s_n - 1)^{-1} \log(\%^{-1})} = 3g^{1/2}$ into the above inequality to obtain

$$P \left[j(n - q_n s_n - 1)^{-1} (n - 2) \left(\frac{\wedge^{(1)'}}{T} \Sigma_{TT}^{(1)-1} \frac{\wedge^{(1)}}{T} \right) \left(\frac{\wedge^{(1)'}}{T} S_{TT}^{(1)-1} \frac{\wedge^{(1)}}{T} \right)^{-1} \right] \leq j \sqrt{16(n - q_n s_n - 1)^{-1} \log(\%^{-1})} = 3g^{1/2} j \mathbb{P} \{ Y_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \} \leq 2\%;$$

which implies that

$$P \left[j(n - q_n s_n - 1)^{-1} (n - 2) \left(\frac{\wedge^{(1)'}}{T} \Sigma_{TT}^{(1)-1} \frac{\wedge^{(1)}}{T} \right) \left(\frac{\wedge^{(1)'}}{T} S_{TT}^{(1)-1} \frac{\wedge^{(1)}}{T} \right)^{-1} \right] \leq j \sqrt{16(n - q_n s_n - 1)^{-1} \log(\%^{-1})} = 3g^{1/2} j \mathbb{P} \{ Y_i = y_i \mathcal{G}_{i=1}^n \setminus \mathcal{M}_n \} \leq (1 - 2\%) \quad (99)$$

Therefore, it can be seen that

$$\begin{aligned} & P \left[j(n - q_n s_n - 1)^{-1} (n - 2) \left(\frac{\wedge^{(1)'}}{T} \Sigma_{TT}^{(1)-1} \frac{\wedge^{(1)}}{T} \right) \left(\frac{\wedge^{(1)'}}{T} S_{TT}^{(1)-1} \frac{\wedge^{(1)}}{T} \right)^{-1} \right] \leq j \sqrt{16(n - q_n s_n - 1)^{-1} \log(\%^{-1})} = 3g^{1/2} \\ & \quad \times \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P \left[j(n - q_n s_n - 1)^{-1} (n - 2) \left(\frac{\wedge^{(1)'}}{T} \Sigma_{TT}^{(1)-1} \frac{\wedge^{(1)}}{T} \right) \left(\frac{\wedge^{(1)'}}{T} S_{TT}^{(1)-1} \frac{\wedge^{(1)}}{T} \right)^{-1} \right] \\ & \quad \leq j \sqrt{16(n - q_n s_n - 1)^{-1} \log(\%^{-1})} = 3g^{1/2} j \mathbb{P} \{ Y_i = y_i \mathcal{G}_{i=1}^n \} = (1 - 2\%) \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P \{ Y_i = y_i \mathcal{G}_{i=1}^n \} \\ & \quad = (1 - 2\%) \sum_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P \{ Y_i = y_i \mathcal{G}_{i=1}^n \} = (1 - 2\%) P(\mathcal{M}_n) \\ & \quad \leq (1 - 2\%) \leq 2 \exp(-n^{-1}) \leq 2 \exp(-n^{-2}); \end{aligned} \quad (100)$$

where the second inequality is by (99), and the last inequality follows from Lemma 3. To this end, based on condition (a), it is straightforward to verify that for any $\% \geq (e^{-n/100}; 1=100)$,

$$\mathcal{M}_{1n}^*(\%) \subseteq \mathcal{M}_{1n}(\%); \quad (101)$$

in which $\mathcal{M}_{1n}^*(\%) = \left[j(n - q_n s_n - 1)^{-1} (n - 2) \left(\frac{\wedge^{(1)'}}{T} \Sigma_{TT}^{(1)-1} \frac{\wedge^{(1)}}{T} \right) \left(\frac{\wedge^{(1)'}}{T} S_{TT}^{(1)-1} \frac{\wedge^{(1)}}{T} \right)^{-1} \right] \leq j \sqrt{16(n - q_n s_n - 1)^{-1} \log(\%^{-1})} = 3g^{1/2}$. Finally, the assertion follows immediately from (100) and (101). \square

Moreover, conditional on $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$, we also note that

$$\begin{aligned} & (n_1^{-1} n_2^{-1} n^2)(q_n s_n = n) + 2(n_1^{-1} n_2^{-1} n^2) f \log(\%^{-1}) = ng \\ & + 2(n_1^{-1} n_2^{-1} n^2)(q_n s_n = n + 2n_1 n_2 n^{-2} \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T})^{1/2} f \log(\%^{-1}) = ng^{1/2} \\ & (400 \binom{-1}{1} \binom{-1}{2})^{1/2} q_n s_n = n + \log(\%^{-1}) = n + f \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T} \log(\%^{-1}) = ng^{1/2} ; \end{aligned}$$

according to the definition of \mathcal{M}_n in Lemma 3. Therefore, based on the above two inequalities, we have

$$\begin{aligned} P^{\wedge \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \wedge \binom{(1)}{T}} & \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T} (400 \binom{-1}{1} \binom{-1}{2})^{1/2} q_n s_n = n + \log(\%^{-1}) = n \\ & + f \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T} \log(\%^{-1}) = ng^{1/2} \quad fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad 1 \quad \% : \end{aligned} \quad (103)$$

Analogously, based on (102) and (8.35) of Lemma 8.1 in Birge (2001), it is obvious that for any $t > 0$,

$$\begin{aligned} P^{\wedge \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \wedge \binom{(1)}{T}} & \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T} (n_1^{-1} n_2^{-1} n^2)(q_n s_n = n) \quad 2(n_1^{-1} n_2^{-1} n^2) \\ & (q_n s_n = n + 2n_1 n_2 n^{-2} \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T})^{1/2} (t=n)^{1/2} \quad fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \\ & \exp(-t) : \end{aligned}$$

We then substitute $t = \log(\%^{-1})$ into the above inequality to obtain

$$\begin{aligned} P^{\wedge \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \wedge \binom{(1)}{T}} & \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T} (n_1^{-1} n_2^{-1} n^2)(q_n s_n = n) \quad 2(n_1^{-1} n_2^{-1} n^2) \\ & (q_n s_n = n + 2n_1 n_2 n^{-2} \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T})^{1/2} f \log(\%^{-1}) = ng^{1/2} \quad fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad 1 \quad \% : \end{aligned}$$

Likewise, we note that conditional on $fY_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$,

$$\begin{aligned} & (n_1^{-1} n_2^{-1} n^2)(q_n s_n = n) \quad 2(n_1^{-1} n_2^{-1} n^2)(q_n s_n = n + 2n_1 n_2 n^{-2} \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T})^{1/2} f \log(\%^{-1}) = ng^{1/2} \\ & (400 \binom{-1}{1} \binom{-1}{2})^{1/2} \log(\%^{-1}) = n + f \binom{(1)'}{T} \Sigma_{TT}^{(1)-1} \binom{(1)}{T} \log(\%^{-1}) = ng^{1/2} ; \end{aligned}$$

We then derive from the above two inequalities that

$$P \left(\bigcap_{T=1}^n \sum_{TT}^{(1)-1} \wedge_T^{(1)} \right) \leq (400 \frac{-1}{1} \frac{-1}{2})^{1/2} \log(\frac{-1}{\%}) = n \\ + f_T \sum_{TT}^{(1)-1} \wedge_T^{(1)} \log(\frac{-1}{\%}) = ng^{1/2} \quad \text{f}Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad \text{1} \quad \%:$$

Together with (103), we arrive at

$$P \left(\mathcal{M}_{2n}(\%) \text{ f}Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad \text{1} \quad \text{2}\% \right) \quad (104)$$

Finally, we have

$$P \text{ f} \mathcal{M}_{2n}(\%) g \quad P \text{ f} \mathcal{M}_{2n}(\%) \setminus \mathcal{M}_n g = \prod_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P \left(\mathcal{M}_{2n}(\%) \text{ f}Y_i = y_i g_{i=1}^n \right) \quad P \left(\text{f}Y_i = y_i g_{i=1}^n \right) \\ = \prod_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} (1 \quad \text{2}\%) \quad P \left(\text{f}Y_i = y_i g_{i=1}^n \right) = (1 \quad \text{2}\%) P(\mathcal{M}_n) \\ = 1 \quad \text{2}\% \quad 2 \exp(-n \frac{-1}{1} = 12) \quad 2 \exp(-n \frac{-1}{2} = 12);$$

where the second inequality is by (104), and the last inequality follows from Lemma 3.

This finishes the proof. \square

Lemma 20. For any $\% \geq 2 (e^{-n/100}; 1=100)$, denote the event $\mathcal{M}_{4n}(\%)$ as

$$\mathcal{M}_{4n}(\%) = \bigcap_{j=1}^{q_n s_n} e_j \Lambda_T^{(1)1/2} \sum_{TT}^{(1)-1} \wedge_T^{(1)} \quad e_j \Lambda_T^{(1)1/2} \sum_{TT}^{(1)-1} \wedge_T^{(1)} \quad (8 \frac{-1}{1} \frac{-1}{2})^{1/2} \\ \text{f} e_j \Lambda_T^{(1)1/2} \sum_{TT}^{(1)-1} \wedge_T^{(1)1/2} e_j g^{1/2} \text{ f} \log(q_n s_n \frac{-1}{\%}) = ng^{1/2} \quad \text{O};$$

where $e_j : j = 1, \dots, q_n s_n$ denotes the standard basis for $\mathbb{R}^{q_n s_n}$. Then we have the following property:

$$P \text{ f} \mathcal{M}_{4n}(\%) g \quad \text{1} \quad \text{2}\% \quad 2 \exp(-n \frac{-1}{1} = 12) \quad 2 \exp(-n \frac{-1}{2} = 12); \quad \text{8}\% \geq 2 (e^{-n/100}; 1=100).$$

Proof of Lemma 20: First of all, we note that conditional on any nonempty $\text{f}Y_i = y_i g_{i=1}^n \setminus$

\mathcal{M}_n

$$\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)} \mathbb{P} Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad N(\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)}; \rho_1^{-1} \rho_2^{-1} \rho \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)1/2}): \quad (105)$$

Moreover, it can be observed that

$$P \mathcal{M}_{4n}(\%) \mathbb{P} Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad \times^{q_n s_n} \quad P \mathcal{M}_{4nj}(\%) \mathbb{P} Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad (q_n s_n - 1);$$

where the events $\mathcal{M}_{4nj}(\%) = \bigcap_j e_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)} \quad e_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)} \quad (8 \quad 1^{-1} \quad 2^{-1})^{1/2}$
 $\mathbb{P} e_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)1/2} e_j g^{1/2} \mathbb{P} \log(q_n s_n \%^{-1}) = \rho g^{1/2}$ for all $j \quad q_n s_n$. Under (105), the concentration inequality entails that for all $j \quad q_n s_n$

$$P \mathcal{M}_{4nj}(\%) \mathbb{P} Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad 1 - 2 \exp(-\log(q_n s_n \%^{-1}) g) = 1 - 2 q_n^{-1} s_n^{-1} \%:$$

Putting the above two inequalities together leads to

$$P \mathcal{M}_{4n}(\%) \mathbb{P} Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad 1 - 2\%: \quad (106)$$

Therefore, we have

$$\begin{aligned} P \mathbb{P} \mathcal{M}_{4n}(\%) g &= P \mathbb{P} \mathcal{M}_{4n}(\%) \setminus \mathcal{M}_n g \\ (1 - 2\%) &\times \\ \{y_i\}_{i=1}^n \in \mathcal{M}_n &P \mathbb{P} Y_i = y_i g_{i=1}^n = (1 - 2\%) P(\mathcal{M}_n) \\ 1 - 2\% &= 2 \exp(-n \cdot 12) = 2 \exp(-n \cdot 2=12); \end{aligned}$$

where the second inequality is by (106), and the last inequality follows from Lemma 3.

This finishes the proof. \square

Lemma 21. For any $\% \geq (e^{-n/100}; 1=100)$, denote the event $\mathcal{M}_{5n}(\%)$ as

$$\begin{aligned} \mathcal{M}_{5n}(\%) &= \bigcap_T \wedge_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\wedge_T^{(1)}) \quad \wedge_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\wedge_T^{(1)}) \\ &\quad (8 \quad 1^{-1} \quad 2^{-1})^{1/2} \frac{1}{\max} (\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)1/2}) \mathbb{P} q_n s_n \log(\%^{-1}) = \rho g^{1/2} : \end{aligned}$$

Then we have the following property:

$$P f M_{5n}(\%) g \quad 1 \quad 2\% \quad 2 \exp(n_1=12) \quad 2 \exp(n_2=12); \quad 8\% \quad 2 (e^{-n/100}; 1=100).$$

Proof of Lemma 21: First of all, we know that conditional on any nonempty $f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$

$$\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)}) f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$$

$$N \nu_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)}), n_1^{-1} n_2^{-1} n f \text{sgn}(\beta_T^{(1)})' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)}) g .$$

Together with the concentration inequality, we conclude that for any $t > 0$

$$P \left(\hat{\nu}_T^{(1)'} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\beta_T^{(1)}) f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \right. \\ \left. 1 \quad 2 \exp \left(-8^{-1} \frac{1}{2} f q_n S_n \max(\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2}) g^{-1} n t^2 \right) : \right.$$

Plugging $t = (8^{-1} \frac{1}{2})^{1/2} \frac{1}{\max(\Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2})} f q_n S_n \log(\frac{1}{\%}) = n g^{1/2}$ into the above inequality yields

$$P \mathcal{M}_{5n}(\%) f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad 1 \quad 2\%: \quad (107)$$

Finally, we have

$$P f M_{5n}(\%) g \quad P f M_{5n}(\%) \setminus \mathcal{M}_n g \\ (1 \quad 2\%) \times_{\{y_i\}_{i=1}^n \in \mathcal{M}_n} P f Y_i = y_i g_{i=1}^n = (1 \quad 2\%) P(\mathcal{M}_n) \\ 1 \quad 2\% \quad 2 \exp(n_1=12) \quad 2 \exp(n_2=12);$$

where the second inequality is by (107), and the last inequality follows from Lemma 3.

This completes the proof. \square

Lemma 22. *Assume the following condition (a):*

$$(a) \quad q_n S_n = o(n).$$

Then there exists universal constants $c_1 > 0$ and $c_2 > 0$ such that:

$$1) P \max_{j \leq q_n s_n} (e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) = (e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) \quad 1$$

$$c_1 q_n s_n = n + \text{fllog}(q_n s_n \log n) = n g^{1/2} \quad 1 \quad c_2 \text{fllog}(n) g^{-1} + \exp(n_1=12) + \exp(n_2=12) .$$

$$2) P \max_{j \leq q_n s_n} (e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) = (e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) \quad 1$$

$$c_1 q_n s_n = n + \text{fllog}(q_n s_n \log n) = n g^{1/2} \quad 1 \quad c_2 \text{fllog}(n) g^{-1} + \exp(n_1=12) + \exp(n_2=12) .$$

$$3) P \text{flsgn} \left(\begin{pmatrix} 1 \\ T \end{pmatrix}' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{pmatrix} 1 \\ T \end{pmatrix} \right) g \right) = \text{flsgn} \left(\begin{pmatrix} 1 \\ T \end{pmatrix}' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{pmatrix} 1 \\ T \end{pmatrix} \right) g \right)$$

$$1 \quad c_1 q_n s_n = n + \text{fllog} \log(n) = n g^{1/2} \quad 1 \quad c_2 \text{fllog}(n) g^{-1} + \exp(n_1=12) + \exp(n_2=12) .$$

$$4) P \text{flsgn} \left(\begin{pmatrix} 1 \\ T \end{pmatrix}' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{pmatrix} 1 \\ T \end{pmatrix} \right) g \right) = \text{flsgn} \left(\begin{pmatrix} 1 \\ T \end{pmatrix}' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{pmatrix} 1 \\ T \end{pmatrix} \right) g \right)$$

$$1 \quad c_1 q_n s_n = n + \text{fllog} \log(n) = n g^{1/2} \quad 1 \quad c_2 \text{fllog}(n) g^{-1} + \exp(n_1=12) + \exp(n_2=12) .$$

Recall that $f e_j : j \leq q_n s_n g$ denotes the standard basis for $\mathbb{R}^{q_n s_n}$.

Proof of Lemma 22: First of all, according to (98), condition (a) and Theorem 3.2.12 in Muirhead (1982), we know that conditional on any nonempty $f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$, and for every $j \leq q_n s_n$,

$$(n-2) (e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) (e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j)^{-1} j f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad \frac{2}{n - q_n s_n - 1} .$$

Together with (A.2) and (A.3) in Johnstone and Lu (2009), it can be deduced that for any $t \in [0; 1-2)$ and for every $j \leq q_n s_n$,

$$P j (n - q_n s_n - 1)^{-1} (n-2) (e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) (e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j)^{-1}$$

$$1 j \quad t j f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad 2 \exp f \quad 3 (n - q_n s_n - 1) t^2 = 16g;$$

which together with condition (a) implies that

$$\begin{aligned}
& P \max_{j \leq q_n s_n} j (\epsilon_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) (\epsilon_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j)^{-1} \\
& \quad 4q_n s_n = n + 2t j f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad 1 \quad 2 \exp f \quad 3(n \quad q_n s_n \quad 1) t^2 = 16g \\
& \quad 1 \quad 2 \exp(\quad n t^2 = 16):
\end{aligned}$$

Together with the union bound inequality, it can be observed that for any $t \geq [0; 1=2)$,

$$\begin{aligned}
& P \max_{j \leq q_n s_n} j (\epsilon_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) (\epsilon_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j)^{-1} \\
& \quad 4q_n s_n = n + 2t j f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \quad 1 \quad 2q_n s_n \exp(\quad n t^2 = 16):
\end{aligned}$$

Subsequently, we substitute $t = \sqrt{16 \log(q_n s_n \log n)} = n g^{1/2}$ into the above inequality to obtain

$$\begin{aligned}
& P \max_{j \leq q_n s_n} j (\epsilon_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) (\epsilon_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j)^{-1} \\
& \quad 4q_n s_n = n + 8 \sqrt{16 \log(q_n s_n \log n)} j f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \\
& \quad 1 \quad 2 \sqrt{16 \log(n)} g^{-1}: \tag{108}
\end{aligned}$$

It then follows that

$$\begin{aligned}
& P \max_{j \leq q_n s_n} (\epsilon_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) (\epsilon_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j)^{-1} \\
& \quad 8 \quad q_n s_n = n + \sqrt{16 \log(q_n s_n \log n)} = n g^{1/2} \\
& \quad \times \\
& \quad \{y_i\}_{i=1}^n \in \mathcal{M}_n \quad P \max_{j \leq q_n s_n} (\epsilon_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) (\epsilon_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) \\
& \quad 1 \quad 8 \quad q_n s_n = n + \sqrt{16 \log(q_n s_n \log n)} = n g^{1/2} \quad f Y_i = y_i g_{i=1}^n \quad P \quad f Y_i = y_i g_{i=1}^n \\
& \quad [1 \quad 2 \sqrt{16 \log(n)} g^{-1}] \quad \times \quad P \quad f Y_i = y_i g_{i=1}^n = [1 \quad 2 \sqrt{16 \log(n)} g^{-1}] P(\mathcal{M}_n) \\
& \quad \{y_i\}_{i=1}^n \in \mathcal{M}_n \\
& \quad 1 \quad 2 \quad \sqrt{16 \log(n)} g^{-1} + \exp(\quad n \quad 1 = 12) + \exp(\quad n \quad 2 = 12) ;
\end{aligned}$$

where the second inequality is by (108), and the last inequality follows from Lemma 3.

Hence, property 1) is justified by the above inequality. To prove property 2), notice that

under the event $\bigcap_{j \leq q_n s_n} (e_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) =$
 $(e_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) \quad 1 \quad 8 \quad q_n s_n = n + f \log(q_n s_n \log n) = n g^{1/2} \quad \circ$, it is straightforward
to verify that

$$\max_{j \leq q_n s_n} (e_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) = (e_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) \quad 1$$

$$2 \max_{j \leq q_n s_n} (e_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) = (e_j \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j) \quad 1 :$$

Putting the above conditions together, we complete the proof of Theorem 3.4. Similar reason can be used to prove Theorem 3.5. \square

where

$$\begin{aligned}\Omega_{1j} &= j e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)} e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)} j; \\ \Omega_{2j} &= j e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)} e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)} j;\end{aligned}$$

Invoking Lemma 20, it can be deduced that there exist universal constants $c_1 > 0$ and $c_2 > 0$ such that

$$\begin{aligned}P \sum_{j=1}^n \Omega_{1j} &\leq c_1 \int \log(q_n s_n \log n) = n g^{1/2} f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{1/2} \\ 1 - c_2 \int \log(n) g^{-1} + \exp(-n_1=12) + \exp(-n_2=12) &:\end{aligned}\tag{110}$$

Regarding the term Ω_{2j} , it can be seen that

$$\Omega_{2j} = f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g \int \Pi_{1j} j (1 + \Pi_{2j}) + (\Omega_{1j} + j e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)} j) \int \Pi_{2j};$$

where

$$\begin{aligned}\Pi_{1j} &= f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)} g f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{-1} \\ &\quad f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \wedge_T^{(1)} g f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{-1}; \\ \Pi_{2j} &= f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{-1} - 1:\end{aligned}$$

For the term Π_{2j} , it follows from Lemma 22 that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that

$$\begin{aligned}P \max_{j \leq q_n s_n} \Pi_{2j} &\leq c_3 [q_n s_n = n + \int \log(q_n s_n \log n) = n g^{1/2}] \\ 1 - c_4 \int \log(n) g^{-1} + \exp(-n_1=12) + \exp(-n_2=12) &:\end{aligned}$$

To this end, based on the above three inequalities, we conclude that there exist universal

constants $c_5 > 0$ and $c_6 > 0$ such that

$$\begin{aligned}
& P \int_{j=1}^{h_{q_n s_n n}} \Omega_{2j} \leq c_5 \prod_{j=1}^h \int \prod_{j=1}^h f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g + [q_n s_n = n + f \log(q_n s_n \log n) = n g^{1/2}] \\
& f \log(q_n s_n \log n) = n g^{1/2} \int f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{1/2} + [q_n s_n = n + f \log(q_n s_n \log n) = n g^{1/2}] \\
& \int e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j \quad \text{O} \quad 1 \leq c_6 f \log(n) g^{-1} + \exp(-n^{-1}) + \exp(-n^{-2}) :
\end{aligned} \tag{111}$$

To bound the term Π_{1j} , for every $j \leq q_n s_n$, we define a $2 \times q_n s_n$ random matrix \hat{M}_j as

$$\hat{M}_j = [\Lambda_T^{(1)1/2} e_j; \hat{\Sigma}_T] \in \mathbb{R}^{2 \times q_n s_n}.$$

Elementary algebra shows that for every $j \leq q_n s_n$,

$$\begin{aligned}
\hat{M}_j \Sigma_{TT}^{(1)-1} \hat{M}_j' &= \begin{bmatrix} e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j & e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T \\ e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T & \hat{\Sigma}_T \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T \end{bmatrix} \in \mathbb{R}^{2 \times 2}; \\
\hat{M}_j \Sigma_{TT}^{(1)-1} \hat{M}_j' &= \begin{bmatrix} e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j & e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T \\ e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T & \hat{\Sigma}_T \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T \end{bmatrix} \in \mathbb{R}^{2 \times 2}.
\end{aligned} \tag{112}$$

Moreover, since $\hat{\Sigma}_T$ is independent of $\Sigma_{TT}^{(1)}$, it can be shown that conditional on any nonempty

$f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{\Sigma}_T g$, and for every $j \leq q_n s_n$,

$$(n-2)(\hat{M}_j \Sigma_{TT}^{(1)-1} \hat{M}_j')^{-1} j f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{\Sigma}_T g \sim \text{Wishart}(n - q_n s_n j (\hat{M}_j \Sigma_{TT}^{(1)-1} \hat{M}_j')^{-1}); \tag{113}$$

using Theorem 3.2.11 in [Muirhead \(1982\)](#). To this end, by combining (112), (113) with

Theorem 3(d) in [Bodnar and Okhrin \(2008\)](#), it is straightforward to reach a conclusion

that for every $j \leq q_n s_n$,

$$f(n - q_n s_n - 3) = \int_j g^{1/2} \Pi_{1j} j f Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{\Sigma}_T g \sim t(n - q_n s_n - 3);$$

where $t(n - q_n s_n - 3)$ represents the student t-distribution with $n - q_n s_n - 3$ degrees of free-

dom, and $\int_j = \int f' \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T g f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{-1} \int f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \hat{\Sigma}_T g^2 f e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{-2}$.

Together with Lemma 20 in [Kolar and Liu \(2015\)](#), it is clear that there exist universal constants $c_7 > 0$ and $c_8 > 0$ such that for every $j \leq q_n s_n$ and for any $t_j \geq 0$,

$$P \prod_{j=1}^{q_n s_n} \left(t_j \mathbb{P}(Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_{TG}) \leq c_7 \exp \left(-c_8 (n - q_n s_n - 3) t_j^2 \right) \right. \\ \left. c_7 \exp \left(-2^{-1} c_8 n \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f}_{TG}^{-1} \hat{f}_{e_j} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g_{t_j}^2 \right) \right);$$

which further implies that

$$P \prod_{j=1}^{q_n s_n} \left(t_j g \mathbb{P}(Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_{TG}) \leq c_7 \exp \left(-2^{-1} c_8 n \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f}_{TG}^{-1} \hat{f}_{e_j} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g_{t_j}^2 \right) \right);$$

By plugging $t_j = c_9 \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f}_{TG}^{1/2} \hat{f}_{e_j} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{-1/2} \lceil \log(q_n s_n \log n) \rceil = n g^{1/2}$ with $c_9 = (2c_8^{-1})^{1/2}$ into the above inequality, it can be obtained that

$$P \prod_{j=1}^{q_n s_n} \left(c_9 \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f}_{TG}^{1/2} \hat{f}_{e_j} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{-1/2} \lceil \log(q_n s_n \log n) \rceil = n g^{1/2} \right. \\ \left. \mathbb{P}(Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \setminus \hat{f}_{TG}) \leq c_7 \lceil \log(n) \rceil g^{-1} \right); \quad (114)$$

It then follows that

$$P \prod_{j=1}^{q_n s_n} \left(c_9 \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f}_{TG}^{1/2} \hat{f}_{e_j} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{-1/2} \lceil \log(q_n s_n \log n) \rceil = n g^{1/2} \right. \\ \left. \times \prod_{\{y_i\}_{i=1}^n \in \mathcal{M}_n, \hat{\nu}_T \in \mathcal{M}_n} P \prod_{j=1}^{q_n s_n} \left(c_9 \hat{f}'_{T \Sigma_{TT}^{(1)-1}} \hat{f}_{TG}^{1/2} \hat{f}_{e_j} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{-1/2} \right. \right. \\ \left. \lceil \log(q_n s_n \log n) \rceil = n g^{1/2} \mathbb{P}(Y_i = y_i g_{i=1}^n \setminus \hat{f}_{TG}) \right) P \mathbb{P}(Y_i = y_i g_{i=1}^n \setminus \hat{f}_{TG}) \\ \left. \leq [1 - c_7 \lceil \log(n) \rceil g^{-1}] \prod_{\{y_i\}_{i=1}^n \in \mathcal{M}_n, \hat{\nu}_T \in \mathcal{M}_n} P \mathbb{P}(Y_i = y_i g_{i=1}^n \setminus \hat{f}_{TG}) = [1 - c_7 \lceil \log(n) \rceil g^{-1}] P(\mathcal{M}_n) \right. \\ \left. \leq 1 - c_{10} \lceil \log(n) \rceil g^{-1} + \exp(-n^{-1}) + \exp(-n^{-2}) \right);$$

for some universal constant $c_{10} > 0$, where the second inequality is by (114). Together with

Lemma 19, it is seen that there exist universal constants $c_{11} > 0$ and $c_{12} > 0$ such that,

$$\begin{aligned}
 P & \prod_{j=1}^{\lfloor \log \log n \rfloor} c_{11} \prod_{j=1}^{\lfloor \log \log n \rfloor} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j g^{-1/2} \log(q_n s_n \log n) = n g^{1/2} \\
 & q_n s_n = n + \log \log(n) = n + [1 + q_n s_n = n + \log \log(n) = n g^{1/2}] \prod_{T'}^{(1)'} \Sigma_{TT}^{(1)-1} \prod_{T'}^{(1)'} g \\
 & + \log \log(n) = n g^{1/2} \prod_{T'}^{(1)'} \Sigma_{TT}^{(1)-1} \prod_{T'}^{(1)'} g^{1/2} \quad \text{O} \\
 & 1 - c_{12} \log(n) g^{-1} + \exp(-n^{-1/2}) + \exp(-n^{-2/2}) :
 \end{aligned}$$

Together with (111), it is clear that there exist universal constants

(b) $c_1 = \min(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2})$ $\max(\Lambda_T^{(1)-1/2} \Sigma_{TT}^{(1)} \Lambda_T^{(1)-1/2})$ c_2 , for some universal constants $0 < c_1 < c_2$.

Then there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that:

$$\begin{aligned} P & \sum_{j=1}^{\lfloor n \rfloor} j e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) j \\ & c_3 \lfloor n \rfloor \log(n) = n g^{1/2} + \lfloor n \rfloor \log \log(n) = n g^{1/2} \\ & + c_3 \sum_{j=1}^{\lfloor n \rfloor} j e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) j \quad \lfloor n \rfloor \log(n) = n g^{1/2} \\ & 1 - c_4 [(\lfloor n \rfloor)^{-1} + \lfloor n \rfloor g^{-1} + \exp(-\lfloor n \rfloor) + \exp(-\lfloor n \rfloor^2)]: \end{aligned}$$

Proof of Lemma 24: First of all, we note that for every $j \leq \lfloor n \rfloor$,

$$j e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) j = \Omega_{1j} + \Omega_{2j}; \quad (115)$$

where

$$\begin{aligned} \Omega_{1j} &= j e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) j; \\ \Omega_{2j} &= j e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \hat{\Lambda}_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) j; \end{aligned}$$

For the term Ω_{1j} , it is apparent to see that for every $j \leq \lfloor n \rfloor$,

$$\Omega_{1j} \leq c_1^{-1} (1 + \Pi_{1j}) \lfloor n \rfloor \Pi_{2j} j + j e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) j \lfloor n \rfloor; \quad (116)$$

where c_1 is defined in condition (b), and

$$\begin{aligned} \Pi_{1j} &= j e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j' g^{-1} \lfloor n \rfloor; \\ \Pi_{2j} &= j e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) j e_j' \Lambda_T^{(1)1/2} S_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j' g^{-1} \\ & \quad + j e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn} \left(\begin{matrix} (1) \\ T \end{matrix} \right) j e_j' \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} e_j' g^{-1}; \end{aligned}$$

To bound the term Π_{1j} , invoking Lemma 22, it can be seen that there exist universal constants $c_3 > 0$ and $c_4 > 0$ such that with probability at least $1 - c_3 \lfloor n \rfloor g^{-1} +$

$\exp(-n^{-1}) + \exp(-n^{-2})]$,

$$\max_{j \leq q_n s_n} \Pi_{1j} \quad c_4 [q_n s_n = n + \lceil \log(q_n s_n \log n) \rceil] : \quad (117)$$

To bound the term Π_{2j} , based on similar argument as in the proof of Lemma 23, it can be shown that there exist universal constants $c_5 > 0$ and $c_6 > 0$ such that conditional on any nonempty $\mathcal{F}Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n$, and for any $t > 0$,

$$P \left[\bigwedge_{j=1}^{q_n s_n} \lceil \Pi_{2j} \rceil \leq t \mid \mathcal{F}Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \right] \geq 1 - c_5 q_n s_n \exp(-c_6 n (q_n s_n)^{-1} t^2 g).$$

By setting $c_7 = c_6^{-1/2}$ and plugging $t = c_7 \lceil q_n s_n \log(q_n s_n \log n) \rceil = n g^{1/2}$ into the above inequality, it can be obtained that

$$P \left[\max_{j \leq q_n s_n} \lceil \Pi_{2j} \rceil \leq c_7 \lceil q_n s_n \log(q_n s_n \log n) \rceil = n g^{1/2} \mid \mathcal{F}Y_i = y_i g_{i=1}^n \setminus \mathcal{M}_n \right] \geq 1 - c_5 \lceil \log(n) \rceil g^{-1}.$$

Together with Lemma 3, there exist universal constants $c_8 > 0$ and $c_9 > 0$ such that with probability at least $1 - c_8 [\lceil \log(n) \rceil g^{-1} + \exp(-n^{-1}) + \exp(-n^{-2})]$,

$$\max_{j \leq q_n s_n} \Pi_{2j} \leq c_9 \lceil q_n s_n \log(q_n s_n \log n) \rceil = n g^{1/2}. \quad (118)$$

By combining (117), (118) with (116), it is seen that there exist universal constants $c_{10} > 0$ and $c_{11} > 0$ such that

$$P \left[\bigwedge_{j=1}^{q_n s_n} \Omega_{1j} \leq c_{10} \lceil q_n s_n \log(q_n s_n \log n) \rceil = n g^{1/2} + c_{10} j e_j \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} \text{sgn}(\hat{\Lambda}_T^{(1)}) \right] \\ \geq 1 - c_{11} [\lceil \log(n) \rceil g^{-1} + \exp(-n^{-1}) + \exp(-n^{-2})] : \quad (119)$$

To bound the term Ω_{2j} , it can be verified that

$$\max_{j \leq q_n s_n} \Omega_{2j} \leq (q_n s_n)^{1/2} k \Lambda_T^{(1)-1/2} \hat{\Lambda}_T^{(1)1/2} \leq k_{\max} \Lambda_T^{(1)1/2} \Sigma_{TT}^{(1)-1} \Lambda_T^{(1)1/2} k_2.$$

Together with Lemma 5 and Lemma 8, it is seen that there exist universal constants $c_{12}, c_{13} > 0$ such that

$$P \max_{j \leq q_n s_n} \Omega_{2j} \leq c_{12} \sqrt{q_n s_n \log(q_n s_n)} = n g^{1/2} \\ 1 - c_{13} [(q_n s_n)^{-1} + \exp(-n_1/12) + \exp(-n_2/12)]:$$

Together with (115) and (119), the assertion holds trivially, which completes the proof. \square

References

- Birge, L. (2001). An alternative point of view on lepski's method. *State of the Art in Probability and Statistics*, 36:113–133.
- Bodnar, T. and Okhrin, Y. (2008). Properties of the singular, inverse and generalized inverse partitioned wishart distributions. *Journal of Multivariate Analysis*, 99(10):2389–2405.
- Johnstone, I. M. and Lu, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *Journal of the American Statistical Association*, 104(486):682–693.
- Kolar, M. and Liu, H. (2015). Optimal feature selection in high-dimensional discriminant analysis. *IEEE Transactions on Information Theory*, 61(2):1063–1083.
- Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley Series in Probability and Statistics. Wiley.
- Ning, Y. and Liu, H. (2017). A general theory of hypothesis tests and confidence regions for sparse high dimensional models. *The Annals of Statistics*, 45(1):158–195.

Shao, J., Wang, Y., Deng, X., and Wang, S. (2011). Sparse linear discriminant analysis by thresholding for high dimensional data. *The Annals of Statistics*, 39(2):1241–1265.